9

Listening strategies in infancy: the roots of music and language development

Sandra E. Trehub and *Laurel J. Trainor*

9.0 INTRODUCTION

There has been growing interest in the perception of complex auditory (non-speech) patterns but surprisingly little concern for the origins of such abilities. Nevertheless, contemporary perspectives in this domain often embody implicit, if not explicit, claims about such origins. For some (e.g. Leek 1987; Watson 1987), pattern perception processes are not stable and 'hardwired' but rather have considerable plasticity, being dependent on the experience, training, and expectations of the listener. Indeed, the perceptual organization of particular patterns is thought to depend on extensive formal or informal training with similar materials, with attention, listening strategy, and task demands playing a prominent role (Espinoza-Varas and Watson 1989). The assumption, moreover, is that the amount and kind of information that listeners can extract from a sound pattern is determined primarily by the relative difficulty of the task and secondarily by features of the pattern itself. Accordingly, a difficult task would lead to a synthetic or global mode of listening (i.e. poor sensitivity for pattern components) in contrast to an easy task, which would engage analytic listening skills (i.e. enhanced resolution of individual components) and make alternative representations available to the listener. What is clear, however, is that prolonged practice is deemed essential for '*fully developed* perceptual representations of patterns' (Espinoza-Varas and Watson 1989, p. 90).

This position implies that auditory pattern perception is almost entirely learned. Conceivably, however, pattern perception could be given in its entirety by nature, in which case the perception of auditory events would be uniform across individuals, with learning involved only in such things as the memory for specific patterns (e.g. individual voices, tunes) as well as arbitrary associations between auditory and non-auditory events (e.g. words and meanings). The views of most researchers tend to fall somewhere between these extremes. For example, Handel (1989), Krumhansl (1992), and Jones (1990) consider innate perceptual processes such as grouping to set the stage for perceiving complex patterns such as speech and musical sequences. Beyond these grouping processes, however, there is much to be learned.

Thinking in sound: the cognitive psychology of human audition, ed. S. McAdams and E. Bigand. Oxford University Press, 1993, pp. 278–327.

According to Krumhansl (1990), adults' perception of pitch relations in music derives from incidental extraction of the regularities in heard music. For Jones (1981, 1982), as well, exposure to music promotes the internalization of prototypes or conventions of a musical culture, with such prototypes generating expectancies that guide the listener's attention. Handel (1989) suggests that innate factors might play a somewhat larger role. For him, heredity supplies the potential perceptual structures that experience 'fine tunes' to the conventions of a particular language or musical system. Just as children acquire their language early, easily, and without formal training, so might they become sensitive to the musical conventions of their culture (Handel 1989, p. 381).

Despite the presumed role of knowledge and experience, few have evaluated these presumptions with novice listeners, notably infants. Such naïve listeners offer unique opportunities for exploring the presence and nature of human pattern processing predispositions as well as the relevance of such predispositions for adult pattern perception. A decade or so of systematic research on the localization of sound patterns in infancy has yielded much valuable information (for reviews, see Clifton (1992), Muir *et al.* (1989)). Unfortunately, there has been considerably less research on other aspects of auditory pattern processing in early life.

To counter this deficiency, we have been pursuing answers to a variety of general questions such as the following. What properties of auditory patterns dominate perception in the early months of life? Do these properties remain influential for more mature listeners? Are the grouping processes that characterize auditory pattern perception (see Bigand, Ch. 8 this volume; Bregman 1990, Ch. 2 this volume) operative in infancy, when instructions to listen synthetically or analytically are necessarily precluded? Although relational processing is the norm for adults' perception of auditory sequences, absolute pitch processing is characteristic of various non-human species such as songbirds and monkeys (D'Amato 1988; Hulse *et al.* 1990). One might ask, then, whether infants are more like human adults in this respect or more like non-human listeners.

Research with adults has revealed that *good* auditory patterns, defined in information-theoretic terms (Garner 1970, 1974) or in terms of conformance to cultural conventions (e.g. Cuddy *et al.* 1981) are perceived and remembered more accurately than poorly structured or *bad* patterns. Is it possible, instead, to define *good* auditory patterns with reference to their relative ease of processing by infant listeners? If so, this would remove the experiential and ethnocentric biases inherent in the notion of a *good* pattern. (Note that aesthetic questions would be irrelevant to this designation of patterns as *good* or *bad*.) Obviously, infants of any age have had the benefit of some listening experience. Nevertheless, patterns that are processed readily in early life could be considered to have special status (i.e. as *good* patterns) and to involve innate attentional predispositions or learning preferences (Locke 1990; Marler 1990). If any such *good* patterns can be identified, one can then ask whether their features are in accord with linguistic (e.g. Kuhl 1986; Liberman *et al.* 1967), psychoacoustic (e.g. Terhardt 1974), music-theoretic (e.g. Butler 1989; Schenker 1954), or mathematical (e.g. Balzano 1982) predictions. One can inquire, as well, whether the perceptual processing strategies that are evident in early life are equally applicable to musical and speech patterns. If so, are there implications of common processing strategies for the acquisition of knowledge about music and language?

9.0.1 Selecting the naïve listener

In principle, the ideal naïve listener is the newborn, who is free from the attentional biases associated with experience, training, and expectations. In practice, however, the newborn is a relatively unco-operative listener by virtue of characteristic drowsiness, a disposition to fuss or cry when awake, and an impoverished response repertoire. For reasons of convenience rather than conviction, we study infants from 6 or 7 months to 10 or 11 months of age. The lower age bound is determined by the onset of a highly reliable and unambiguous response (i.e. turning head and eyes toward a sound source) to salient changes in a sound pattern (Eilers *et al.* 1977; Kuhl 1985; Trehub *et al.* 1984). The upper age bound is determined by the onset of walking, which is generally accompanied by a reluctance to remain seated for the duration of a test session (10–20 minutes).

It would not be surprising for the pattern processing abilities of infants in this age range to have undergone some modification as a result of early exposure to particular patterns. Nevertheless, it is reasonable to consider these abilities as reflecting pattern processing predispositions. In other words, whether infants are predisposed to perceive patterns in particular ways or are predisposed to learn to perceive them in such ways is immaterial.

9.0.2 Selecting the stimuli

In our research on auditory pattern perception, we use musical or music-like patterns for a number of reasons. Music, 'a peculiarly human adaptation to life' (Slobin and Titon 1984, p. 9), is found in every culture. Although few individuals in Western society perform in public contexts, many sing or hum in the shower or elsewhere, and all listen to music, by choice (e.g. concerts, stereo, radio) or otherwise (e.g. commercial background music). What is not generally known, however, is that contemporary Western notions of musical talent, giftedness, and music as art are unusual (Walker 1987, 1990). Historically and cross-culturally, it has been more common for music to be integrated into various facets of work and play, with all community members participating fully (e.g. Bebey 1969). Moreover, there are many widely held beliefs or myths about music including its influence on physical and mental states, its healing powers, and supernatural origins (Walker 1990). The continuing influence of such belief systems can be seen in contemporary uses of music in promoting religious, political, and commercial goals, parental concerns about so-called 'decadent' music, and young children's greater compliance with sung as opposed to spoken requests. The ubiquity of music and its presumed power raise questions about biological significance (Granit 1977; Lerdahl and Jackendoff 1983) for which there are no ready answers. What is clear is that music is prevalent cross-culturally, occupying an important role in daily life.

A further impetus to the use of musical patterns stems from recent progress in the understanding of music perception by musically trained and untrained adults (Deutsch 1982; Dowling and Harwood 1986; Howell *et al.* 1985; Krumhansl 1990; Sloboda 1985). The presumption is that untrained adults, by virtue of informal but extended exposure to the music of their culture, have developed implicit but highly elaborated schemas for music processing. Those with formal training, on the other hand, have more explicit knowledge of musical structure and are capable of more analytic listening. Is it reasonable to assume, however, that the music processing strategies of untrained adults are due entirely to musical exposure, with no carry-over of primitive processing strategies from early life? Terhardt (1987) suggests, instead, that composers, in creating music, intuitively capitalize on universal principles of auditory perception. It would hardly be surprising, then, for some of these principles to be operative in infancy.

Finally, many have drawn attention to numerous parallels between speech and music (e.g. Handel 1989; Lerdahl and Jackendoff 1983; Sloboda 1985), some even suggesting the possibility of innate perceptual principles for organizing speech and musical input (Krumhansl 1992; Lerdahl and Jackendoff 1983). Handel (1989) has applied some of Hockett's (1963) design features of language to music, as well. These include the features of discreteness (i.e. messages in speech or language being constructed from a limited set of units, whether phonemes or scale notes), openness (i.e. an infinite number of possible linguistic utterances or musical passages), and duality of structure (i.e. meaningless units such as phonemes or notes combined to form meaningful units such as words or musical phrases).

In short, it is not unreasonable to expect that the study of musical pattern perception in infancy will yield valuable information about auditory pattern processing in general. In this spirit, we have used sequences of tones (mostly sine waves) which, despite their harmonic impoverishment, are readily recognizable as melodies. This contrasts with the synthesized speech stimuli in speech perception experiments, some of which are unrecognizable as speech to adult listeners (e.g. Diehl *et al.* 1981). In fact, speech synthesis by rule in conjunction with the elimination of all but a single discriminative cue (the typical procedure for speech perception experiments) can result in very unnatural-sounding stimuli.

9.0.3 The test procedure

We attempt to overload the information-processing capacities of infants by presenting them with five- to ten-note sequences or melodies. Assuming that infants are unable to retain the entire melody, then what they do retain can inform us about their processing strategies. Would they encode and retain an incomplete but precise description such as the absolute pitches of the first or last few notes? Such a local-processing strategy would be uncharacteristic of adults but consistent nevertheless with the preferred strategy of songbirds (Hulse *et al.* 1990) and monkeys (D'Amato 1988). A global-processing alternative might involve configural information about pitch relations, temporal relations, or both. Some researchers suggest that local processing has priority over global processing in infant perception (e.g. Aslin and Smith 1988) whereas others suggest the reverse (e.g. Morrongiello 1988).

We gather indirect evidence for the use of such strategies by means of an operant discrimination procedure. The infant sits on the parent's lap in one corner of a sound-attenuating booth facing an experimenter, who maintains the infant's attention by manipulating puppets (see Fig. 9.1). The standard melody or tone sequence is presented repeatedly from a loudspeaker 45° to the infant's left. The experimenter continually records (via push buttons) when the infant is looking directly ahead (i.e. ready for a test trial) and when the infant turns toward the loudspeaker (i.e. a correct or false-positive response). The parent and the experimenter wear headphones with masking sounds so that both remain uninformed about the patterns being presented. Test trials, which are presented only when the infant is looking directly ahead, are of two types: *change* trials in which a comparison sequence, which embodies subtle or substantial changes, replaces the standard sequence for one or two repetitions, and *no-change* trials in which the comparison sequence is the same as the standard. A

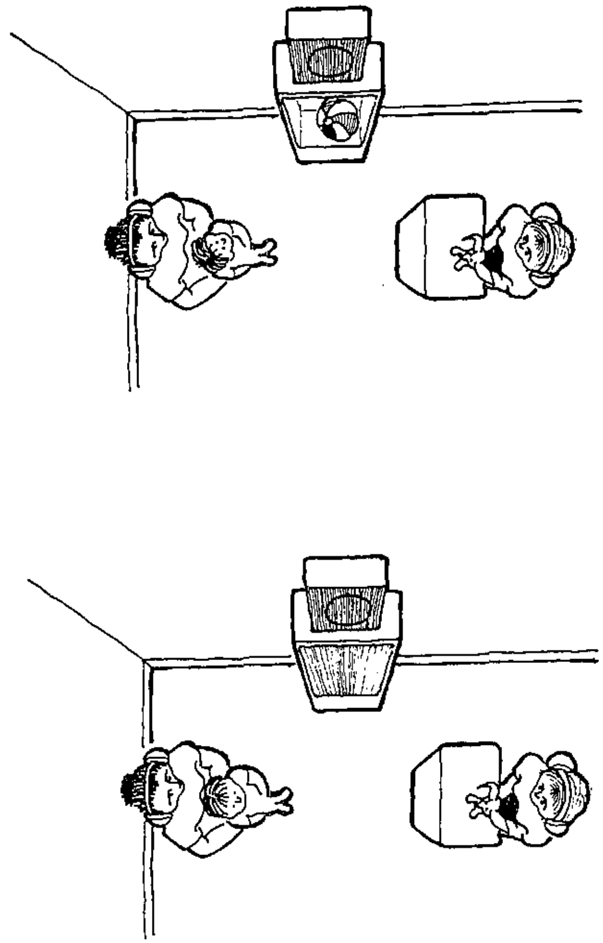

Fig. 9.1 The test situation. The infant sits on the parent's lap facing the experimenter. Sounds are presented repeatedly through the loudspeaker to the infant's left. Turns to the speaker on change trials lead to the presentation of an animated toy near the speaker (From Trehub (1990).)

head turn (45° or more) toward the loudspeaker within 4 s of the sound change results in the illumination and activation of one of four animated toys near the loudspeaker. Turns at other times (i.e. false-positive responses) are unreinforced.

The difficulty of the task can be varied by manipulating the length and complexity of the sequences, the size of the retention interval (i.e. time between standard and comparison sequences), and the degree of deviation from standard to comparison sequences. It is also possible to introduce variations in stimulus parameters across repetitions of the patterns. For example, the repetitions of the standard and comparison patterns could be presented in pitch transposition (i.e. at different pitch levels), precluding the use of absolute pitch cues in the discrimination task. The patterns could also be presented at different tempos or rates, precluding the use of absolute duration cues.

In an initial training phase, infants must meet a training criterion of four consecutive correct responses to a more prominent sound change than that used in the test phase. The subsequent test phase consists of approximately 30 test trials (more or less in some experiments), with a random ordering of 15 change and 15 no-change trials. If infants respond significantly more often on change than on no-change trials, this indicates that they can detect the change.

We make the leap from discrimination data to processing strategy in the following way. If infants are using a particular processing strategy such as encoding the exact beginnings of melodies, then certain changes would evoke a response (e.g. new initial pitches) and others would not (e.g. new final pitches). Likewise, a global-processing strategy would be reflected in infants' failure to respond to certain discriminable changes such as new pitches if the pitch configuration or contour remains the same (e.g. transpositions), or new note durations if the relative note durations (i.e. rhythmic structure) remain the same.

9.1 GROUPING

Perceptual grouping is thought to operate preattentively (Neisser 1967), creating units that optimize cognitive processing (Bregman 1981, 1990). Cues for promoting grouping such as the relative proximity of elements and their similarity were described by the Gestalt psychologists many years ago but their role in auditory contexts has only become clear in recent years (Bigand, Ch. 8 this volume; Bregman 1990, Ch. 2 this volume; Deutsch 1982).

The presumption is that at least some grouping processes are primitive (Bregman 1990) or informationally encapsulated (Fodor 1983), being relatively insensitive to the perceiver's experience. This raises the possibility that these processes would be operative early in life, even in infancy. On the other hand, some grouping phenomena are clearly dependent on experience, that is, they are schema-based (Bregman 1990). For example, the organization of the speech stream into words creates the illusion of pauses between words (Studdert-Kennedy 1975), an illusion that is absent when the sounds belong to a foreign language. In general, however, we tend to group subsets of elements within auditory sequences, even when the elements are totally uniform (Fraisse 1982). Such perceptual grouping has consequences for judgments of duration (Bolton 1894; Woodrow 1909) as well as duration discrimination (Fitzgibbons *et al.* 1974).

If infants grouped auditory sequences in similar ways, we might expect them to have greater difficulty detecting a silent interval or pause *between* groups of sounds compared with an identical pause *within* a group of sounds. In a series of studies (Thorpe and Trehub 1989; Thorpe *et al.* 1988), we assessed the ability of infants 6–9 months of age to detect small temporal changes in patterns of six temporally equidistant tones. The patterns were structured so that the first three tones, which were identical, differed from the last three (also identical to each other) in fundamental frequency (pitch), spectral structure (i.e. waveform: sawtooth vs. sine waves), or intensity. The schematic structure can be represented as XXXOOO. The comparison sequences incorporated increments to the silent interval between the third and fourth tone (XXX OOO), a between-group change, or between the fourth and fifth tone (XXXO OO), a within-group change. Infants detected the temporal alterations when they occurred within a group but not between groups, implying that they had grouped the original pattern of temporally equidistant tones. As a result, the change that altered or disrupted the structure of the standard pattern (from a 3–3 to a 4–2 grouping) was noticeable but not the change that conserved the structure (3–3 in both cases).

Adults tested with the identical sequences but with smaller increment values and a more conventional psychophysical task (two-alternative, forced-choice task) showed a similar pattern of performance (Thorpe 1985). We call this phenomenon the *duration illusion* (Thorpe and Trehub 1989), referring to the inaccuracy of perceiving the duration of silent intervals *between* groups compared to identical silent intervals *within* a group of sounds. This illusion of pauses between sound groups is analogous to the illusion of pauses between words. Discontinuities in pitch, timbre, or loudness typically signal important aspects of pattern structure, triggering the duration illusion, which guides primitive processes of segmentation or parsing. Greater discontinuity or change in any of these parameters leads to greater inaccuracy in detecting the between-group intervals. In effect, listeners override the available temporal information (e.g. temporally equidistant tones), imposing a temporal structure that corresponds to configural aspects (e.g. pattern of pitches) of the pattern.

Another segmentation process that has received considerably more experimental attention with adults is *auditory stream segregation* (Bregman 1978, 1981) or *auditory scene analysis* (Bregman 1990, Ch. 2 this volume). Rapidly repeating sequences of discrete sounds are grouped or segregated on the basis of similarities in frequency, spectral envelope, or other salient properties, leading to the perception of two or more parallel sequences or streams. One consequence of stream segregation is that listeners are unable to track the order of elements across streams but can readily do so within a single stream (e.g. Bregman and Campbell 1971). Bregman (1990) envisions the operation of unlearned as well as learned constraints on auditory scene analysis, with the former reflecting universal regularities in the environment.

Demany (1982) investigated the phenomenon with infants, comparing 7- to 15-week-old infants' discrimination of the temporal order of two sequences of tones. For adult listeners, one of the sequences would remain within a single auditory stream whereas the other would segregate into different streams. Infants showed a pattern of performance similar to that of adults, detecting a change in the temporal order of tones within the hypothesized stream but failing to detect this change across streams. The implication, then, is that stream segregation, like the duration illusion, reflects primitive rather than schema-based parsing.

Our findings and those of Demany (1982) indicate that infants extract the conventional temporal organization of simple auditory sequences. Moreover, infants can also differentiate patterns with contrasting rhythmic structure (Demany *et al.* 1977; Chang and Trehub 1977*a*; Morrongiello 1984). Adults go beyond this, however, generalizing the temporal or rhythmic structure of sequences across variations in tempo or rate. Would infants do likewise? We evaluated their ability to differentiate three-tone sequences with 1–2 (X XX) or 2–1 (XX X) structure and four-tone sequences with 2–2 (XX XX) or 3–1 (XXX X) structure in the context of tempo variations (Trehub and Thorpe 1989). In the training phase, infants were presented with a three- or four-tone pattern at a uniform tempo but at five different pitch levels (changing randomly from one repetition to the next) and were rewarded for responding to the contrasting three- or four-tone pattern with contrasting structure presented at the same tempo (and variable pitch level). In the test phase, we added five (discriminable) tempo variations to the five pitch variations and rewarded infants only for responding to changes in temporal structure. Infants discriminated between patterns with contrasting temporal structure in the context of variations in tempo and frequency (see Fig. 9.2), indicating that they perceived the similarities in rhythmic structure across these variations. Nevertheless, infants found the task difficult, as reflected in their low level of performance and high false-alarm rate. Such a high false-alarm rate has also characterized infants' performance on other difficult tasks with the same procedure (Thorpe and Trehub 1989). In any case, their performance provided evidence of relational processing in the temporal domain.

This does not appear to be a uniquely human feat. Relational temporal processing has been observed in starlings, which are capable of transposing a rhythmic discrimination across changes in tempo (Hulse *et al.* 1984) and in pigeons, which can respond to the longer of two temporal intervals (Fetterman 1987). Nevertheless, such temporal processing is central to

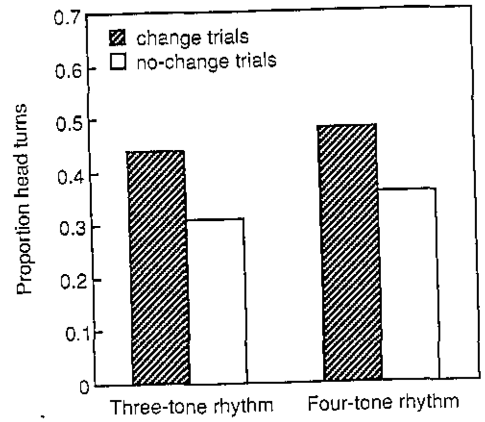

Fig. 9.2 Mean proportion of head turns on change and no-change trials for the three-tone and four-tone rhythms. (Data from Trehub and Thorpe (1989).)

the perception of speech. A verbal message maintains its integrity over changes in rate just as a tune does. There is evidence, moreover, that infants (Eimas and Miller 1980), like adults (Miller and Liberman 1979; Summerfield 1981), make appropriate perceptual compensations for speaking rate.

In a novel approach to the study of grouping in musical patterns, Krumhansl and Jusczyk (1990) explored 4- and 6-month-old infants' sensitivity to phrase structure in music. They presented infants with two distorted versions of Mozart minuets and attempted to establish infants' preference for one over the other. In one version, they added a one-second pause to the end of each phrase, distorting the overall temporal patterning but maintaining the temporal integrity of each musical phrase. In the other, they added the same number of pauses but these were inserted within phrases. This manipulation altered the phrase structure as well as the overall temporal structure of the musical passages. For each infant, the minuets with between-phrase pauses were presented from a loudspeaker on one side and those with within-phrase pauses were presented from the other side. The time spent looking at each loudspeaker during sound presentation was calculated over a series of trials. Infants looked significantly longer at the loudspeaker with intact phrases than at the one with distorted phrases, indicating their 'preference' for the former.

Krumhansl and Jusczyk's (1990) analysis of musical features of the minuets indicated that drops in pitch height and increased duration of the last melody note characterized phrase endings. They suggest that these features signal phrase structure or perceptual units for infant listeners and may account for their preference for intact phrases.

This preference procedure has also been used to explore infants' temporal processing of speech sequences. In these studies (Hirsh-Pasek *et al.* 1987; Kemler Nelson *et al.* 1989), infants heard extended samples of speech in which pauses had been added either between clauses or within clauses. As was the case with music, infants looked longer in the direction of speech with intact clauses (i.e. between-clause pauses) than in the direction of speech with temporally distorted clauses. The implication is that, like musical phrases, spoken clauses are perceptual units, even for prelinguistic infants. Interestingly, this 'preference' for intact clauses was limited to speech directed to infants, being absent in speech directed to adults.

9.2 MELODY PERCEPTION: PITCH CONTOUR

A cursory outline of the melody perception skills of untrained adult listeners can provide a useful backdrop to the consideration of infant skills. It is well known that adults' recognition of melodies is independent of specific pitch levels, being dependent instead on the relations among component pitches. With familiar melodies, the pattern of *intervals* is relevant (Attneave and Olson 1971; Dowling and Fujitani 1971), intervals referring to the precise pitch relations between adjacent notes. Thus transpositions, which have different component notes but an identical pattern of intervals, are perceived as equivalent to one another (Attneave and Olson 1971). With unfamiliar melodies, configural information about pitch is prominent, notably the *melodic contour* or global pattern of changes in pitch direction (up/down/same) (Bartlett and Dowling 1980; Dowling 1978). In addition, unfamiliar melodies that conform to familiar musical principles are encoded in greater detail, remembered more readily, and are also preferred compared with those that deviate from such principles (Cuddy *et al.* 1981; Francès 1988; Krumhansl and Keil 1982; Krumhansl *et al.* 1982; Watkins 1985; Zenatti 1969).

Studies of melodic processing in non-human species have revealed qualitative differences. To the extent that melody discrimination has been demonstrated in monkeys (D'Amato and Salmon 1984), starlings (Hulse and Cynx 1986), and budgerigars (Dooling *et al.* 1987), local (absolute) pitch cues rather than global pattern cues (e.g. contour) have been implicated. By contrast, the perception of single speech sounds by various non-human species, including budgerigars (Dooling *et al.* 1990), chinchillas (Kuhl and Miller 1975), and monkeys (Kuhl and Padden 1982) has revealed remarkable parallels between human and non-human listeners.

Does this imply that the perception of single sounds, including speech sounds, capitalizes on principles common to avian and mammalian species, whereas human adult-like perception of sound sequences depends on experience and the resultant meaningfulness of such patterns? If we adopt a conventional definition of auditory patterns, referring to 'sounds characterized by a perceptual impression that is global . . . elicited by the overall spectral or temporal form' (Espinoza-Varas and Watson 1989, p. 70), then we would have to conclude that the various non-human species that have been studied to date are unable to perceive pitch patterns. What about human infants?

In a series of studies, we focused on three types of melodic information: *contour* (i.e. pitch configuration or up/down/same pattern of pitch change), *interval* (i.e. frequency ratio of successive notes or pitch distance in semitones), and *absolute pitch* (i.e. exact pitch level). Our goal was to determine whether infants' mental representation of melodies was based on absolute pitches, exact intervals, or contour. In one study (Trehub *et al.* 1984), we tested infants on their discrimination of various changes to a six-tone melody including *transpositions* (i.e. absolute pitch changes, but same intervals and contour), *contour-preserving* changes (i.e. absolute pitch and interval changes, but same contour), and *contour-violating* changes (e.g. interval and contour changes, but same pitches reordered). To preclude the use of obvious cues, the comparison melodies had the same initial and final note as the standard melody except for the transpositions, which necessitated different notes (see Fig. 9.3). For adult listeners, the transposed comparison would seem most similar to the original melody (essentially equivalent to it) and the contour-violating comparison most dissimilar. The contour-preserving changes would be intermediate on the similarity–dissimilarity continuum. In the context of a difficult task such as a single presentation of the standard pattern and a long retention interval, adults might have comparable difficulty detecting contour-preserving changes and transpositions.

Overall, infants' performance was consistent with that expected for adults. Infants performed well on the contour changes but not on the transpositions and contour-preserving (i.e. interval) changes (see Fig. 9.4). In fact, they responded to the latter two types of changed melody as though they were further repetitions of the standard melody. Essentially, their performance was consistent with a global processing strategy, with information about pitch contour dominating at the expense of absolute pitch and interval information. For example, the transpositions had the same contour but contained six new notes and infants seemed to respond on the basis of the unchanged contour. Thus, they treated the transposed melodies as familiar rather than as novel or changed patterns. It is possible, however, that infants responded on the basis of the interval pattern, which was also invariant across transpositions. This seems unlikely

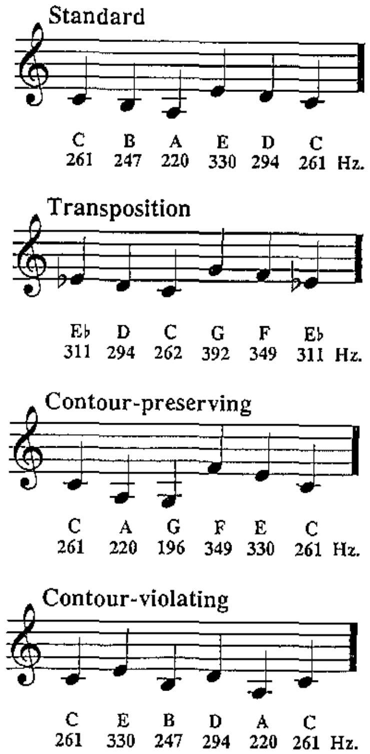

Fig. 9.3 One of the standard melodies and three types of changes to be detected: transposition (same contour, same intervals, different pitches); contour-preserving change (same contour, different intervals, different pitches); contour-violating change (different contour, different intervals, same pitches). (Stimuli from Trehub *et al.* (1984).)

because infants also treated melodies with new intervals but the same contour as though they were familiar patterns. In fact, they responded identically to transpositions and to contour-preserving changes, implying contour-influenced performance in both cases. Moreover, their ease of detecting contour changes was not limited to situations with multiple directional changes in pitch movement. Rather, infants also responded readily to contour-violating comparisons that retained five of the six original notes in their original order but had one note that changed the contour (Trehub *et al.* 1985).

Infants accomplished these contour discriminations with a fixed standard or background melody (i.e. standard and comparison sequences presented at one pitch level) so that absolute pitch cues were available. When we eliminated the potential use of absolute cues (Trehub *et al.* 1987) by presenting a simple standard pattern (up–down contour) repeating

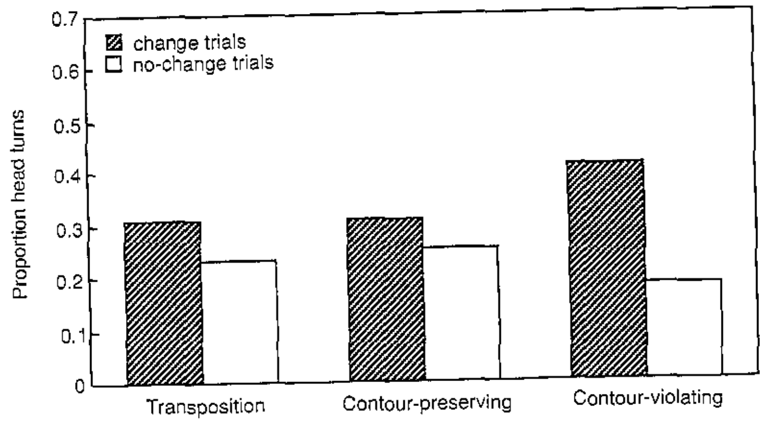

Fig 9.4 Mean proportion of head turns on change and no-change trials for the transposition, contour-preserving, and contour-violating changes. (Data from Trehub *et al.* (1984).)

either in transposition or with changing intervals (but same contour) and a contour-violating comparison (down–up–down contour) with reordered pitches (see Fig. 9.5), infants readily detected the contour changes (see Fig. 9.6). For the most part, they ignored the local pitch changes so long as the contour was retained, responding when the contour changed. Their relational processing strategy was qualitatively similar to adults' characteristic strategy with unfamiliar melodies but distinctly different from the local pitch strategy of non-human species. Note, also, that infants' performance on the contour discrimination tasks was considerably better than their performance on the temporal processing tasks with the same (head-turning) procedure (Thorpe and Trehub 1989; Thorpe *et al.* 1988; Trehub and Thorpe 1989).

These findings suggest that pitch-contour processing may function as an important perceptual organizational device for infants, guiding their parsing or segmentation of complex auditory patterns. Infants have other, perhaps less potent, parsing devices at their disposal including temporal grouping (by pitch, timbre, or loudness) and rhythmic patterning. These organizational processes would be useful not only for processing musical sequences but also for speech sequences, particularly prosodic aspects of speech.

Prosody is thought to be replete with cues to important linguistic units or boundaries (Morgan *et al.* 1987), cues that are enhanced in the unique prosody of speech directed to infants. Indeed, the hallmark of infant-directed speech is its distinctive and perhaps universal pitch contours (Fernald *et al.* 1989; Stern *et al.* 1983). Such speech also embodies higher pitch (by three or four semitones), an increased pitch range, smoother

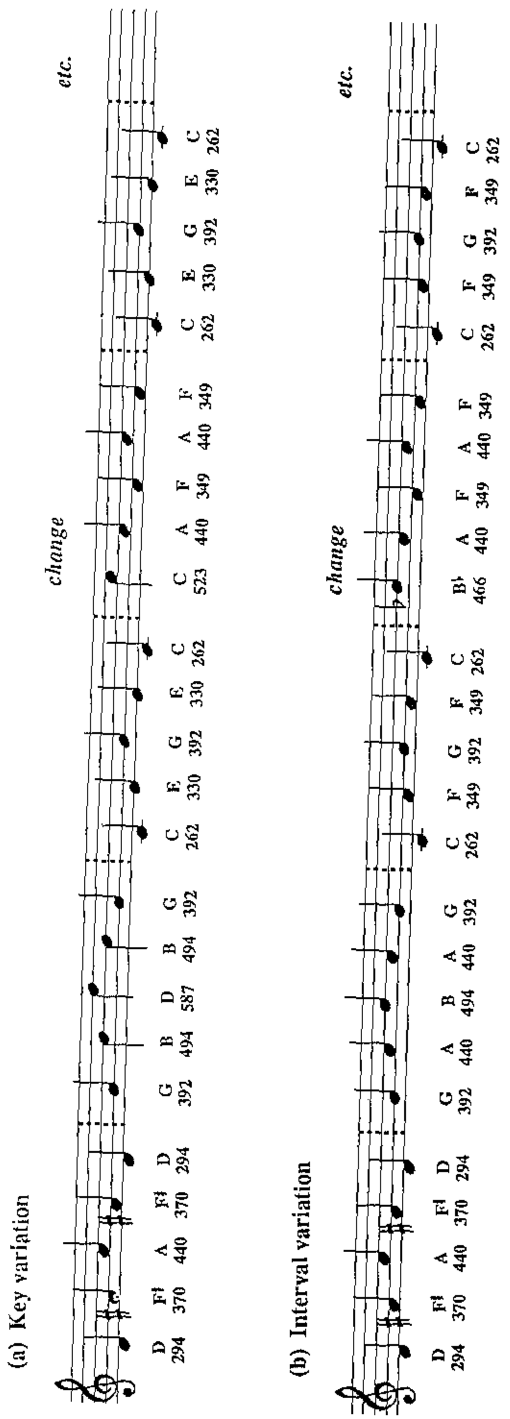

Fig. 9.5 Examples of repeating stimuli for the key-variation (a) and interval-variation (b) conditions. Each example depicts three successive presentations of the standard melody followed by a change and a subsequent return to the standard melody. (Stimuli from Trehub *et al.* (1987).)

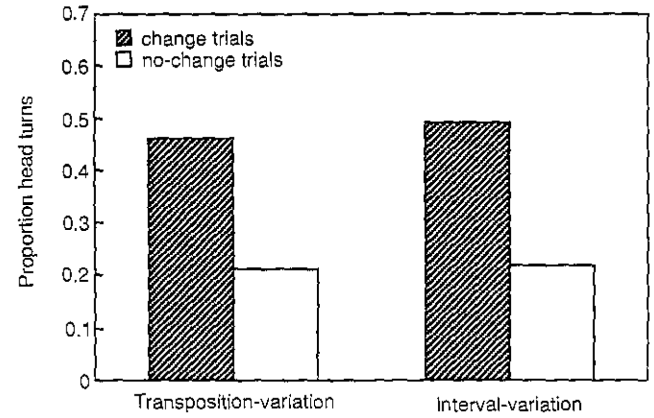

Fig. 9.6 Mean proportion of head turns on change and no-change trials for the transposition-variation and interval-variation conditions. (Data from Trehub *et al.* (1987).)

pitch transitions, simpler pitch contours, slower tempo, more regular rhythms, shorter utterances, and greater repetitiveness compared with adult-directed speech (Fernald and Simon 1984; Fernald *et al.* 1989; Papoušek and Papoušek 1981; Papoušek *et al.* 1985; Stern *et al.* 1983). These various qualities lend unity or coherence to the pitch contours of infant-directed speech, setting them apart from the complex contours of adult-directed speech (Stern *et al.* 1982). In fact, these distinctive prosodic forms are rarely produced by the same individuals in non-maternal contexts (Fernald and Simon 1984).

The characteristic features of infant-directed speech have been designated musical by a number of researchers (e.g. Fernald 1989; Papoušek and Papoušek 1981). The typical pitch contours are unidirectional (rising or falling) but at times bell-shaped (rise–fall, fall–rise) and these are repeated over and over with altered lexical or segmental content and varying tempo. Vowels are usually extended, as in song, and many utterances are contentless from a semantic perspective, consisting of a monosyllable (consonant–vowel or vowel alone) with the vowel stretched out over an expanded pitch contour (Fernald and Simon 1984; Papoušek and Papoušek 1981).

The regular rhythms and slow tempo of infant-directed speech (Beebe *et al.* 1985) also confer a musical quality to such speech. In fact, rhythmic adult-directed speech is atypical, being confined, for the most part, to ritualized or pathological contexts (Jaffe *et al.* 1979). It is also atypical of children's speech, except for ritual chants (Hargreaves 1986; Moorhead and Pond 1978).

There are indications that specific pitch contours are linked universally to particular caretaking contexts (Fernald *et al.* 1989). For example, rising or bell-shaped contours are used to gain and maintain infant attention (Papoušek and Papoušek 1981), falling contours with a narrow pitch range to soothe and hasten sleep (Papoušek and Papoušek 1981), and more variable contours to heighten positive affect (Stern *et al.* 1982). The tempo and rhythm of maternal utterances are also tuned to the infant's presumed needs. Thus slow, rhythmic utterances are provided for attentive infants, increased tempo for inattentive infants, variable rhythms and tempo for fussy infants, and gradually decreasing tempo for infants progressing toward sleep (Papoušek and Papoušek 1981). In sum, maternal vocal stimulation seems to be tuned intuitively to infants' processing abilities in general (Trehub 1989, 1990) and to their fluctuating state in particular.

This care-giving speech register has a dramatic impact on infant listeners. From as young as two days of age, they listen preferentially to infant-directed over adult-directed speech (Cooper and Aslin 1990; Fernald 1985; Werker and McLeod 1989), even to synthesized sine-wave renditions of infant-directed over adult-directed pitch contours (Fernald and Kuhl 1987). Infants also exhibit more positive affect when listening to infant-directed than to adult-directed speech (Werker and McLeod 1989). The universality of a small set of contours tied to specific caretaking contexts and the impact of such contours on infant attention and arousal have prompted some authors to consider their biological basis (Papoušek and Papoušek 1987) and communicative significance (Fernald 1989, 1992).

The pitch contour may well be an important unit of speech processing for prelinguistic listeners, helping to define the boundaries of infant-directed utterances (Trehub 1990). We know that young infants are sensitive to prosodic markers of clauses, phrases, and words in native and non-native languages (Hirsh-Pasek *et al.* 1987; Jusczyk, personal communication; Kemler Nelson *et al.* 1989), perhaps using temporal as well as pitch contour cues to delineate such units. We also know that infants are sensitive to musical phrase structure, with one of the likely cues to phrase boundaries being a falling pitch contour (Krumhansl and Jusczyk 1990).

Although infant-directed speech is music-like in a number of respects, it still differs considerably from the music with which most of us are familiar. If we consider infant-directed song, however, more commonalities might emerge between speech and music. Recently, we unobtrusively recorded mothers of young infants as they informally sang a song of their choice in two contexts, one with their infant present, the other when the mothers were alone (Trehub *et al.* 1993*b*). Most of the chosen materials were play songs, which were appropriate to the presence of awake and lively infants. A tape recording of pairs of identical excerpts from both conditions (infant present, infant absent) was presented

to adult listeners, who were required to identify the infant-directed excerpt. Because the musical materials were the same, identification would have to be based on aspects of the singers' performance. Listeners were highly accurate (about 90 per cent correct) in identifying the infant-directed singing of North American mothers. They were also accurate in differentiating *actual* infant-directed singing from *simulated* infant-directed singing in another group of mothers (Trehub *et al.* 1993*a*). They were less accurate (about 60 per cent correct) but still above chance in judging comparable excerpts from a foreign language (Hindi) and musical system (Indian) (Trehub *et al.* 1993*b*). Not surprisingly, listeners of Indian origin were more accurate than native North American listeners with the Hindi songs.

One musical genre, the lullaby, is sung by care-givers throughout the world to soothe infants and induce sleep (Brakeley 1950; Cass-Beggs and Cass-Beggs 1969). There are indications, moreover, that lullabies are structurally and functionally distinct from other songs just as infant-directed speech is structurally and functionally distinct from other speech. For example, lullabies of the Cuna Indians of Panama embody more liberal textual and melodic improvisation than do other songs (McCosker 1974). The text, which includes word reduplication, sequence repetition, and very common words, is incorporated into repetitive rhythmic patterns. Unlike other Cuna songs, Cuna lullabies are indefinitely long, continuing until the infant listener is quieted or asleep. On the melodic front, Cuna lullabies have a narrow pitch range and repeating contours, much like soothing infant-directed speech, which is characterized by low falling contours, a narrow pitch range, and a gentle tone (Fernald and Simon 1984; Papoušek and Papoušek 1981). Vietnamese (Cong-Huyen-Ton-Nu 1979), Afghani Hazara (Sakata 1987), and North American Indian (Sands and Sekaquaptewa 1978) lullabies, among others, also have smooth, repeating contours, a narrow pitch range, extended vowels, and repetitive rhythms.

In the Hazara culture of central Afghanistan, mothers deliberately alter words to produce mellifluous sounds, with sound effects rather than meaning in mind (Sakata 1987). Indeed, the use of humming and stereotyped syllables such as *loo-loo*, *lulla*, *ninna*, *bo-bo*, and *do-do* is cross-culturally pervasive (Brakeley 1950; Brown 1980). Although lullabies seem to have a distinctive form and style within a culture and some common features across cultures, it is not clear whether any feature or quality of lullabies is universal.

Adults seem to be unaware of the full range of their vocal adjustments in infant-directed speech (Fernald and Simon 1984; Papoušek and Papoušek 1987) but the prosody of such speech is clearly recognizable to adult listeners. In fact, adults easily identify several patterns of maternal prosody including those associated with approval, prohibition, soothing, game-playing, and attempts to capture infant attention (Fernald 1989). Just as universal prosodic features are uncommon in adult-directed speech but are frequent and recognizable in infant-directed speech, so similar universals may emerge in infant-directed music despite their reported absence in other types of music (Harwood 1976).

Recently, we collected a large sample of foreign lullabies and comparison songs (non-lullabies) from diverse cultures and geographic regions (Unyk *et al.* 1992). We prepared a tape in which each of 30 lullabies was paired with a song that was similar in tempo and overall musical style. Adult listeners were required to judge which of the two songs in each pair was a lullaby or song for infants (Trehub *et al.* 1992). They identified the lullabies significantly better than chance (about 63 per cent correct) and performance was unrelated to listeners' level of musical training or their familiarity with any musical culture. When the songs were low-pass filtered (i.e. information below 500 Hz retained) to remove the lullaby lyrics, which had potential cues such as word or syllable repetition, extended vowels, and onomatopoeia, the lullabies were still identifiable. Even the elimination of residual voice quality cues by synthesis of the melody line (with piano timbre) resulted in performance that was highly correlated with performance on the original materials. This suggests that there must be melodic as well as word cues to the identity of lullabies. There are indications, moreover, that simple, descending contours play a role in adults' identification of lullabies (Unyk *et al.* 1992). In short, infant-directed music, whether lullabies or play songs, are identifiable to adult listeners.

9.3 MUSICAL STRUCTURE: DEFINING *GOOD* PATTERNS

Most scholars in the domain of music perception have assumed that specialized schemas for processing the structure of music are developed and progressively refined through exposure (Bharucha 1987; Jones 1982; Krumhansl 1983; Zenatti 1969), although the relevant mechanisms are largely unknown. One possibility is that listeners' mental representations of pitch structure are based on the relative frequency of occurrence of various features in the music of their culture (Krumhansl 1990). If this were the case, then listeners exposed to different musical systems would acquire different musical schemas just as listeners exposed to different languages acquire different speech and language schemas.

Comparisons of various languages have revealed commonalities (but also differences) in the inventory of sound types and in the location of category boundaries for speech (Maddieson 1984). At the same time, sounds that occur frequently across languages (e.g. *ma/pa*) tend to be mastered earlier than those that are relatively infrequent (e.g. *la/ra*) (Jakobson 1968; Locke 1983). This is not the case for the relative frequency

of occurrence of sounds within a language. For example, the initial sounds of *this* and *thing* occur frequently in English but they are relatively infrequent across languages and are late acquisitions for English-speaking children. One might argue, then, that the selection of phonemes and phonemic boundaries has been guided, to some extent, by their relative ease of processing (Burnham *et al.* 1991; Maddieson 1984; Stevens and Keyser 1989). Indeed, linguists have made considerable progress in defining featural constraints on the selection of candidate speech sounds (e.g. Comrie 1981).

Is it also the case that different musical systems reflect pattern-processing dispositions that go beyond the perceptual organizational principles outlined earlier? One near-universal musical feature is octave-equivalence (Dowling and Harwood 1986), whereby notes an octave apart (having a ratio of 2:1) are considered to be equivalent in some way, even by infants (Demany and Armand 1984). In the music of many cultures, they are even given the same name. Another universal feature is the scaling of musical pitch perception on the basis of log frequency rather than linear frequency. It is likely that these common features reflect processing constraints of the auditory system.

At the level of pattern structure, however, commonalities are more difficult to find. Most musical systems seem to be based on scales, a scale dividing the octave into a small number of discrete pitches (usually 5 to 7). The size of the pitch set likely reflects limits on the chunks of information that can be maintained in working memory (Dowling 1978), thereby facilitating the analysis and coding of relations among pitches. In most musical systems, moreover, the absolute pitch of notes is less important than their relative pitch so that a melody remains essentially the same regardless of pitch level.

Beyond these general features, however, there are few identifiable constraints on pitch relations in scale structures. Indeed, ethnomusicologists are pessimistic about the prospect of finding further universals (Harwood 1976). Nevertheless, the recognition of lullabies across cultures (Trehub *et al.* 1992) is suggestive of possible universal features of musical form beyond those identified to date (Trehub and Unyk, 1992). Such universal features might be evident in musical patterns that have special status for infant listeners (i.e. *good* patterns as defined here). With this long-range goal in mind, we have been attempting to identify the features of *good* musical patterns for infants. As noted, our use of the term *good* does not imply aesthetic or value judgements. Rather, our concern is with auditory patterns that can be processed more readily than others by human listeners with limited exposure to any musical system. Presumably, this would reveal the pattern-processing proclivities of naïve listeners and provide clues to potential musical universals.

In principle, the choice of musical materials could be virtually unlimited with infant listeners. In practice, however, most of the relevant research with adults has been based on Western music. Fortunately, there is a rich music-theoretical tradition in the West on which perceptual hypotheses can be based. Accordingly, we have begun our search for *good* patterns by examining infants' sensitivity to various structural features of conventional Western music. Our reasoning is that features to which infants are sensitive reflect constituent features of inherently *good* musical patterns in the sense that they require relatively limited exposure for their mastery.

It is likely that many *good* musical features and *good* musical patterns are not represented in conventional Western music and, therefore, have been outside the scope of most psychological research to date. A pan-cultural definition of *good* patterns is obviously desirable but this must await research with foreign musical materials which, in turn, depends on foreign informants who can vouch for the appropriateness of the stimuli.

An understanding of potentially *good* musical patterns requires some background information about Western tonal music. The octave in Western music is divided into 12 equally spaced notes (on a logarithmic frequency scale) that form the chromatic scale. This division is repeated in successive octaves. (For a more comprehensive description of the structure of Western tonal music see Bigand, Ch. 8 this volume.) For convenience, we can number the 12 notes of the chromatic scale 0 to 11. It is not this equal-interval scale but rather the major scale, an unequal-interval subset of notes from the chromatic scale, that is most commonly used in Western musical composition (see Fig. 9.7). The successive intervals between notes of the major scale are as follows: tone (i.e. two semitones), tone, semitone, tone, tone, tone, semitone. For example, notes 0, 2, 4, 5, 7, 9, and 11 of the chromatic scale form a major scale, as do notes 4, 6, 8, 9, 11, 1, and 3. In musical composition, moreover, different notes of the scale have different functions (Piston 1969). For example, the first note or *tonic* is considered the most stable, with melodies typically ending on this note. (For a discussion of the relative stability of notes and chords, see Bharucha (1984)). The fifth note or *dominant* is important harmonically, with a chord based on this note generating the expectation of resolution to a chord based on the tonic note. The interval between the tonic and dominant notes is called a *perfect fifth* and consists of seven semitones. Melodically, the notes of the *tonic triad*, that is, the first, third, and fifth notes, are considered most stable in their mental representation (Dowling and Harwood 1986). Musically untrained adults have tacit knowledge of the different functions of various scale notes, as reflected in their ratings of how well each note fits into a major-scale context (Krumhansl 1990).

The tonic triad of a major scale is one instance of a *major triad*. In Western music theory, the major triad (intervals: four semitones, three semitones;

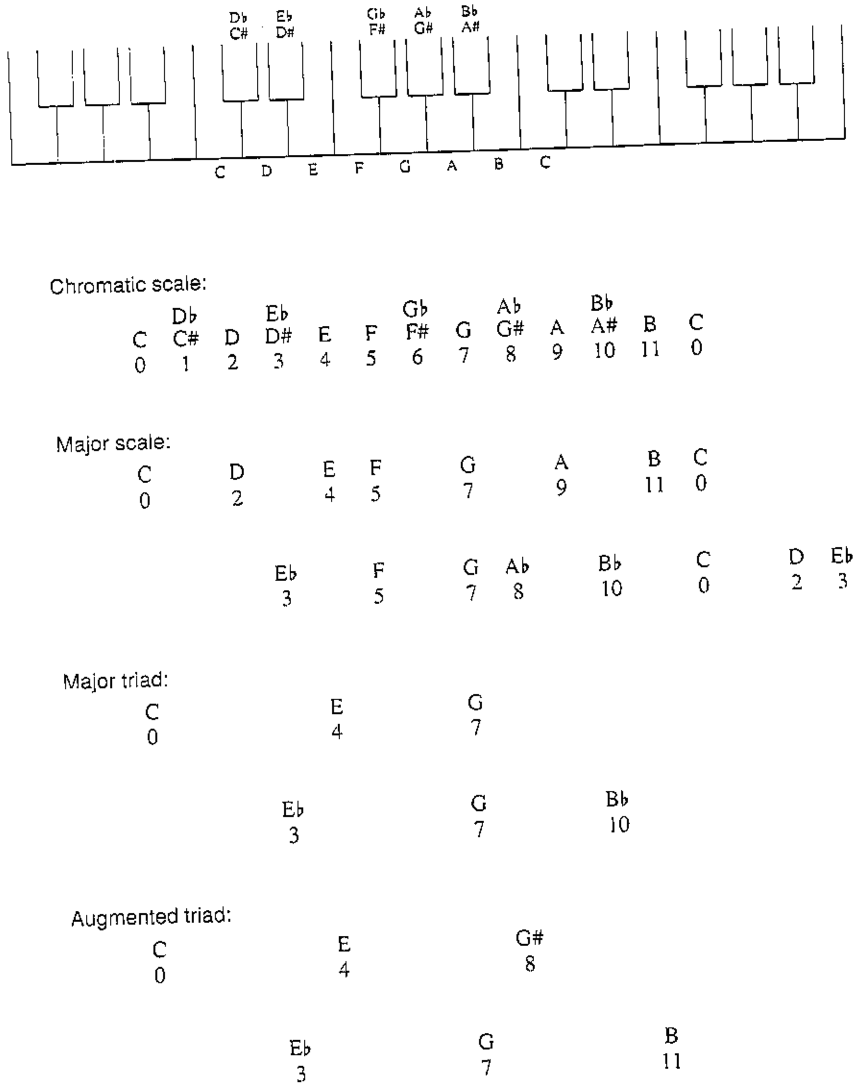

Fig. 9.7 Some fundamentals of Western music structure: the chromatic scale, consisting of successive semitones, two major scales, illustrating the interval pattern 2 2 1 2 2 2 1 semitones; two major triads and two augmented trials, illustrating the interval patterns 4 3 and 4 4 semitones, respectively.

e.g. notes 0, 4, 7) is considered to be a consonant, stable, and perhaps prototypical form (Schenker 1954). Raising the third note of a major triad by one semitone (intervals: four semitones, four semitones; e.g. notes 0, 4, 8) results in an *augmented triad*, which is considered to be dissonant and unstable. It is not known, however, whether the major triad is an intrinsically *good* form or whether it achieves this status only after schemas for Western music are developed.

We addressed this question by testing infants on their ability to distinguish between the major and augmented triads (Cohen *et al.* 1987) (see Fig. 9.8). In one condition, a major triad melody was presented as the standard (e.g., notes 0, 4, 7, 4, 0; or C E G E C in the key of C major), that is, it was repeated over and over in transposition, with successive repetitions beginning on different pitches. From time to time, the note highest in pitch was raised by a semitone (G became G#), forming an augmented triad. The infant's task was to respond to this interval (i.e. contour-preserving) change by turning towards the loudspeaker. In another condition, the standard and comparison melodies were reversed so that the augmented triad melody (C E G#E C) now served as standard and the major as comparison. The infant's task in this case was to detect a downward semitone change (G# became G).

If infants engaged their usual contour-processing strategies, as in previous studies (e.g. Chang and Trehub 1977*b*; Trehub *et al.* 1984, 1987), they would necessarily fail on both conditions of this interval discrimination task. In fact, their performance was asymmetrical, being above chance levels when the major triad (a conventional Western pattern) served as the standard but at chance when the augmented triad (an unconventional Western pattern) was the standard (see Fig. 9.9). It is important to note, however, that performance on this interval discrimination task was much poorer than performance on the contour discrimination tasks outlined earlier. It would seem, then, that contour discrimination is relatively easy for infants whereas interval discrimination, although possible in some situations, is considerably more difficult, as it is with adults (Dowling 1978). Nevertheless, one can still consider the major triad to be a *good* pattern for infant listeners because it permits them to go beyond their usual contour-processing strategy to engage a specifically musical or interval-processing strategy. By contrast, the augmented triad is not a *good* pattern in that infants remain limited to their general-purpose strategy of contour processing.

The asymmetric performance on the major-triad/augmented-triad comparison has numerous parallels in adult perception. For example, adults find it easier to detect changes to conventionally structured melodic (Bharucha 1984; Francès 1988), rhythmic (Bharucha and Pryor 1986), and linguistic (Bharucha *et al.* 1985) patterns than to less conventionally structured patterns, even when changes to the latter result in *good*

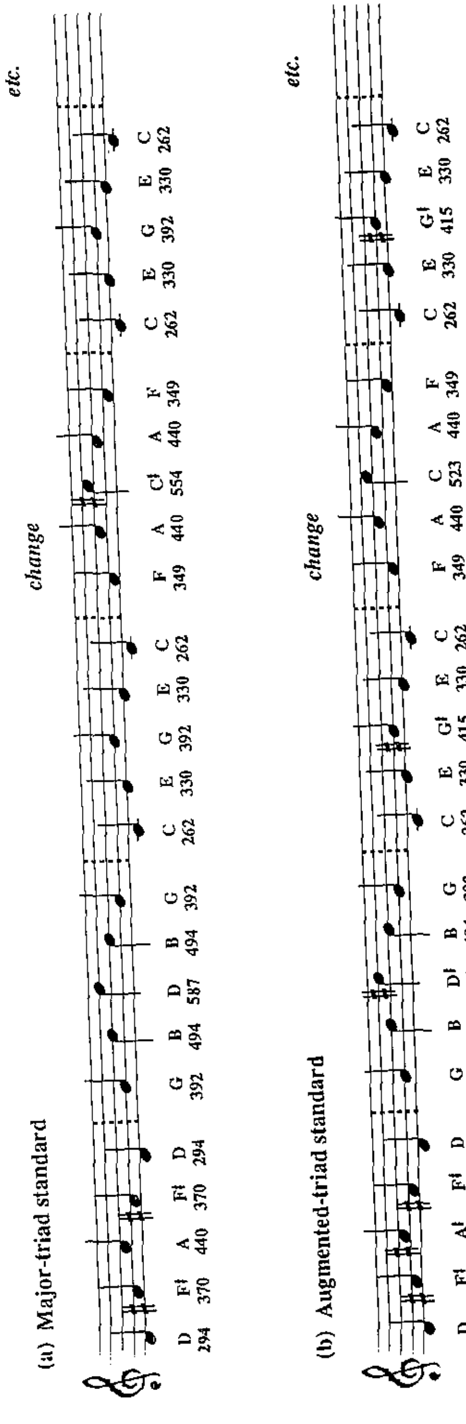

Fig. 9.8 Examples of repeating stimuli with (a) the major-triad melody as standard and the augmented-triad melody as change and (b) the reverse. Each example depicts three successive standard melodies followed by a change and a subsequent return to the standard melody. (Stimuli from Trainor (1991).)

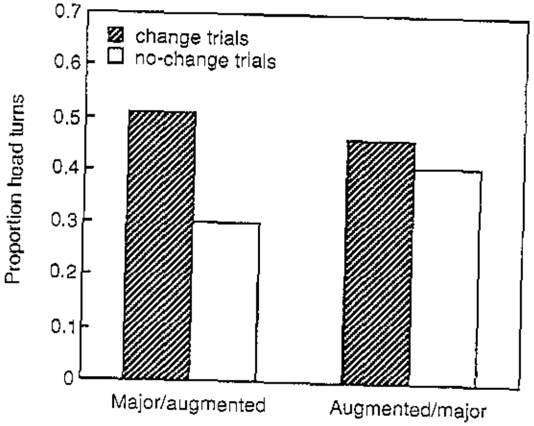

Fig 9.9 Mean proportion of head turns on change and no-change trials with the major-triad and augmented-triad melody serving as standard in one condition and as change in the other. (Data from Trainor (1991).)

patterns. Presumably, the mental representation of a *good* pattern is stable or coherent, thereby facilitating the detection of subtle changes.

There are several possible explanations for the structural significance of the major triad for infant listeners. The major scale is considered to embody special mathematical properties or structural elegance (Balzano 1980, 1982) and the major triad, by virtue of its centrality to major scale structure, would share such properties. According to this perspective, patterns exemplifying major scale structure would be intrinsically easy to process. From a psychoacoustic perspective, intervals approximating simple frequency ratios are considered more consonant than those with more complex ratios (see Burns and Ward 1982; Rakowski 1990). According to this line of reasoning, the notes of the major triad, which approximate simple ratios (4:5:6), would be more consonant than those of the augmented triad, with its complex ratios (16:20:25). In addition, the outer notes of the major triad are related by the perfect fifth interval, which is prominent in the overtone series. This interval is prevalent in naturally occurring sounds, including the simultaneous components of vowel sounds. As a result, ecological considerations may also be implicated in the special status of the major triad.

There is another sense in which the major and augmented triad melodies of these studies were simple, being bilaterally symmetrical and having only three different notes (i.e. the last two notes of these five-note patterns were mirror images of the first two notes). Because of the potential significance of infants' enhanced processing for a conventional Western pattern, it was important to replicate this finding. It was also the case

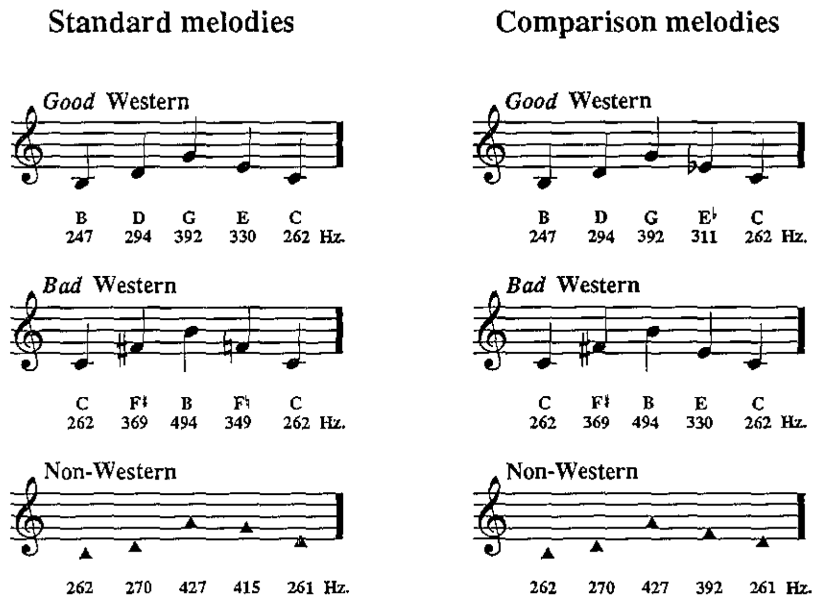

Fig. 9.10 Standard and comparison melodies in *good* Western, *bad* Western, and non-Western conditions. (Note that the non-Western melodies cannot be written in conventional Western notation.) (Stimuli from Trehub *et al.* (1990).)

that the major/augmented comparisons in this study confounded the structure of the standard pattern (major or augmented triad) with upward or downward changes of a semitone. It is possible, then, that enhanced performance with the major triad as standard was attributable to the upward pitch change rather than its structural superiority. To resolve the issue, we evaluated upward and downward changes to major and augmented melodies (Trainor 1991) and still found enhanced processing for major triad melodies by infants as young as six months of age.

We continued to explore the issue of enhanced processing for conventional Western structure by using somewhat more complex, nonsymmetrical patterns consisting of five different pitches (Trehub *et al.* 1990). We presented infant listeners with three types of standard pattern (see Fig. 9.10), the comparison patterns all having a downward change of a semitone. One pattern was a conventionally structured Western melody that was based on the major triad but did not begin on the tonic (B_3 D_4 G_4 E_4 C_4). Another was less conventionally structured (C_4 $F\#_4$ B_4 F_4 $C\#_4$), its notes derived from the chromatic division of the octave but not from any major scale. The third melody was essentially non-Western in that some of its intervals were not based on semitones or their multiples. In all three conditions, the task consisted of detecting

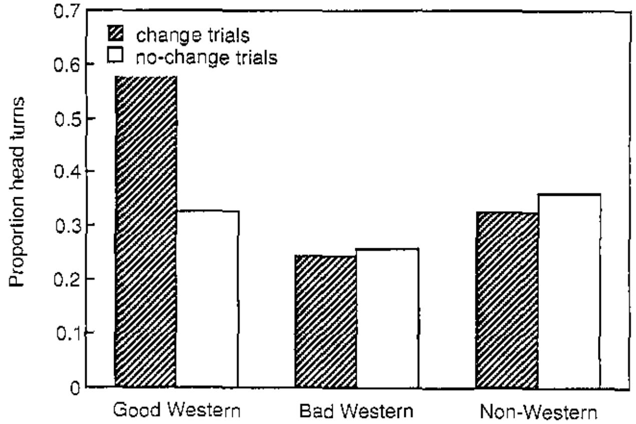

Fig. 9.11 Mean proportion of head turns on change and no-change trials for the *good* Western, *bad* Western, and non-Western melodies. (Data from Trehub *et al.* (1990).)

a change in the fourth note of the melody. We replicated the previous findings of enhanced processing for the major triad in that infants performed above chance levels only on the conventional Western melody (see Fig. 9.11). Naturally, there are differences among the three patterns that go beyond the conventionality of their structure (e.g. interval sizes). Nevertheless, taken together with the previously reported studies of major and augmented patterns, the findings are consistent with the notion that conventional Western melodies confer processing advantages on infants as young as six months of age.

The notion that some objects or patterns serve as cognitive reference points for other exemplars is not new to cognitive psychology (Rosch 1975) although most research of this nature has focused on the visual modality. Such prototypical exemplars are more readily encoded and remembered by adults than are atypical exemplars (Mervis and Rosch 1981; Rosch 1975), as is the case for prototypical musical patterns (e.g. Bharucha 1984; Francès 1988). Even three- and six-month-old infants more readily categorize prototypical visual patterns compared with atypical patterns (Younger and Gottlieb 1988). Moreover, infants respond preferentially to attractive (i.e. as rated by adults) or prototypical faces (Langlois *et al.* 1987; Langlois and Roggman 1990) despite their limited experience with different faces.

In the auditory domain, there are suggestions that some single speech sounds, notably vowels, are rated as better exemplars of their phonetic category than are other instances (Grieser and Kuhl 1989; Kuhl 1991; Miller and Volaitis 1989). Moreover, the rated typicality or *goodness* of a vowel sound exerts substantial effects on perception both for adults and

for six-month-old infants (Grieser and Kuhl 1989; Kuhl 1991). Unfortunately, notions of prototypicality have not been evaluated for speech sound combinations or for speech sequences. It is possible, however, that certain intonation patterns or pitch contours, notably those found in infant-directed speech, might be basic forms from which the more complex contours of adult-directed speech are derived.

9.4 GOOD CONTEXTS: KEY RELATIONS

Context effects are pervasive in audition as they are in vision. For example, the cues for individual speech sounds vary considerably as a function of the preceding and following sounds (e.g. Dorman and Raphael 1980) or the visible articulatory gestures (e.g. McGurk and MacDonald 1976). Similar context effects are evident in music. For example, the major triad provides a superior melodic context for detecting semitone changes than does the augmented triad or other unconventionally structured melodies (e.g. Cohen *et al.* 1987).

Context effects can operate at different levels of generality. The aforementioned context effects concerned the relations among elements of a pattern, whether the sounds of a word or the notes of a melody. One can consider more global or general contexts for individual sounds or sound units such as the syntactic or semantic relations within and between utterances or the pitch relations between musical patterns or phrases. Having established that infants were sensitive to the context exemplified by relations within a melody, we sought to establish whether they were also sensitive to the more general context exemplified by relations *between* melodies.

One important feature of Western scale structure is that a unique major scale can be formed on each of the 12 notes of the chromatic scale, that is, with each of the 12 chromatic notes serving as tonic. In other words, the particular notes of the major scale with tonic 0 are different from the notes of the major scale starting on any other chromatic note. Each scale is considered to be closely related to two other scales, differing from these by only one note (see Fig. 9.12). These highly related scales have tonic notes that are related by the interval of the perfect fifth, which means that the tonic note of one scale of a highly related pair is the dominant of the other. Musical compositions in the Western tonal idiom often modulate or change from one key to another, although they generally end in the key in which they began. The most common modulation, however, is to the key whose tonic is the dominant of the original key. Western adults have internalized these relations, as exemplified by the so-called *key-distance effect* (Bartlett and Dowling 1980; Cuddy *et al.*

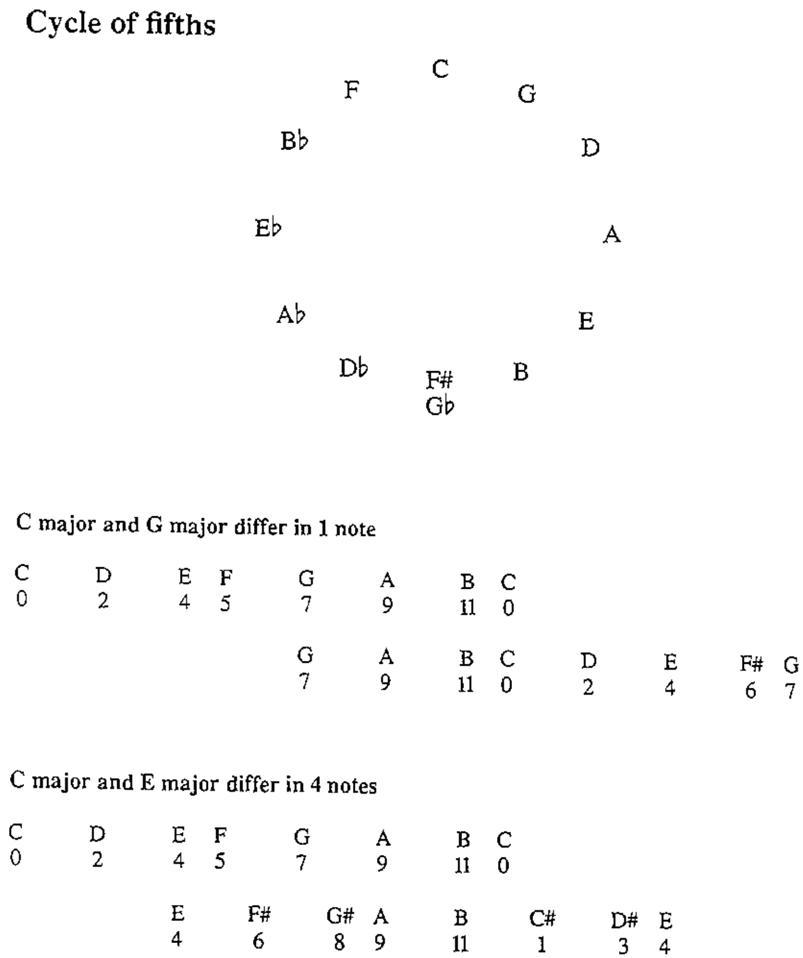

Fig. 9.12 The cycle of fifths; the distance between any two keys on the cycle of fifths is directly proportional to their relatedness.

1979; Krumhansl *et al.* 1982). For example, Cuddy *et al.* (1979) found that adults' discrimination of melodic changes was most accurate when the keys of standard and comparison melodies were related by a perfect fifth. Does this reflect adults' long-term exposure to Western music or does it arise from primitive processing predispositions? If infants' enhanced processing for the major triad arises from the presence of perfect fifth intervals (i.e. the relation between its outer notes), then keys related by perfect fifths might promote comparable enhancement in melodic processing.

Accordingly, we evaluated infants on their detection of a semitone

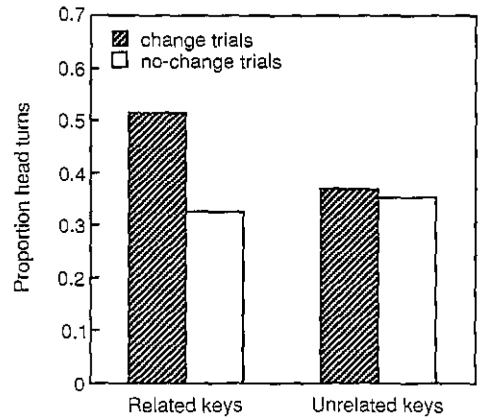

Fig. 9.13 Mean proportion of head turns on change and no-change trials, averaged across major- and augmented-triad melodies, for related and unrelated keys. (Data from Trainor (1991).)

change (upward) in the simple, symmetric major triad melody or augmented triad melody used previously (Trainor and Trehub 1993). In the near-key condition, successive repetitions of the standard and comparison melodies were in closely related keys. The set of keys included B♭, F, C, G, and D, and adjacent melodies were in adjacent keys (in either direction). In the far-key condition, melodies were presented in one of three distantly related keys, C, E, and G# major, again with successive melodies in different keys. Any two of these three keys have only three notes in common (i.e. they differ in four notes). Infants performed better on the major and augmented triad melodies when successive melodies were in related keys (see Fig. 9.13). Thus, interval processing in infancy seems to be enhanced when melodies embody special structures such as the major triad and when musical contexts embody special relations such as those reflected in key relatedness. In short, infants are sensitive not only to the structure of musical patterns but to the macrostructure of pattern repetitions.

It is perhaps significant that the tasks in which infants displayed enhanced processing involved the interval of the perfect fifth. As noted, the outer notes of the major triad form a perfect fifth as do the corresponding notes of melodies transposed to closely related keys. One important but unresolved issue is whether such enhanced processing stems from the simple integer ratio associated with the perfect fifth, that of 3:2. In this regard, it is interesting that individuals from widely different cultures often sing in parallel fifths as well as octaves, thinking that they are singing in unison (Kolinski 1967).

9.5 GOOD SCALES

Recently, there has been increasing interest in the structural characteristics of *good* musical scales, prompted in part by theoretical interest in pattern processing and in part by composers attempting to create new scales suitable for musical composition. Different musical systems seem to use different arrays of intervals, making it possible, in principle, to evaluate the properties of particular scales. In practice, however, most research has been limited to Western scales. Among the findings of such research is that adults have tacit knowledge of major scale structure, as reflected in their ratings of the degree of fit of in-key and out-of-key notes in a major-key context (Krumhansl 1990).

There are claims, as well, that the major scale is special or unique. One approach emphasizes its unique psychoacoustic properties (Burns and Ward 1982; Helmholtz 1954; Krumhansl 1987). With equal-tempered tuning (i.e. all semitones of the chromatic scale equal), the ratios between the fundamental frequency of important scale notes approximate, but do not equal, simple integer ratios, which are considered ideal. For example, notes an octave apart stand in an approximate 2:1 frequency ratio; notes a perfect fifth apart (i.e. seven semitones; e.g. C G) form an approximate 3:2 ratio; notes a perfect fourth apart (five semitones; e.g. C F) an approximate 4:3 ratio; notes a major third apart (i.e. four semitones; e.g. C E) an approximate 5:4 ratio; and so on. Historically, tuning systems embodied these exact ratios but such systems were abandoned because it was only possible for one key to be perfectly in tune at a time, thereby precluding effective key modulation in complex compositions. In line with the notion that small integer ratios are ideal, it is often claimed that singers and players of variable-tuning instruments adjust intervals away from equal temperament toward exact small integer ratios. However, analyses of musical performances are inconsistent with this claim (see Burns and Ward 1982; Rakowski 1990).

Another approach to the uniqueness of the major scale emphasizes mathematical properties of the group of scale notes (Balzano 1980, 1982) independent of psychoacoustic properties. One such property is uniqueness. A scale possesses the uniqueness property if the vector of intervals from each note to every other note in the set is unique. For example, given the set {0 2 4 5 7 9 11}, the vector of intervals formed from the third note (4) is {4–4, 5–4, 7–4, 9–4, 11–4, 12–4, 14–4} or {0 1 3 5 7 8 10}. This pattern of intervals is different from the pattern based on every other note of this scale. On the other hand, all vectors of intervals of the whole-tone scale are the same: {0 2 4 6 8 10}. In this sense, the whole-tone scale is considered to have minimal uniqueness. It is the uniqueness property that allows different notes to take on different functional relations with other notes, facilitating, for example, the tonal hierarchy (Krumhansl 1990).

By contrast, the whole-tone scale has no differentiation of function because all notes are related in the same way. An empirical question, then, is whether these mathematical properties have perceptual consequences, that is, whether melodies based on the major scale are processed more easily than those based on the whole-tone scale.

We are currently evaluating this question by testing infants and adults on their ability to detect a semitone change in a melody based on the major scale (B_3 D_4 G_4 E_4 C_4) or on the whole-tone scale ($B\flat_3$ D_4 $G\#_4$ E_4 C_4). Adults (non-musicians) performed significantly better on the major melody than on the whole-tone melody. Infants showed the same pattern of performance, detecting the change to the major melody but not to the whole-tone melody. This suggests that major scale structure may be more readily encoded than whole-tone scale structure, even by naïve listeners. Despite the failure of cross-cultural studies to uncover universal constraints on intervals in musical scale structure, some scales seem to be intrinsically better than others. Whether this claim is limited to Western infants and adults remains to be determined.

We are conducting further studies with different major and whole-tone melodies and with other scales, both foreign and artificial (Krumhansl 1987; Mathews *et al.* 1987). In particular, the *modal* scales are of interest. For example, the notes 0, 2, 4, 5, 7, 9, 11 form a major scale with tonic 0. These same notes, however, also form a Phrygian modal scale, 4, 5, 7, 9, 11, 0, 2 with tonic 4. Although these scales use the same set of pitches, they are organized around a different tonic so that melodies based on such scales would differ. For example, the melodies would tend to begin and end on different notes and have different statistical properties (i.e. frequency of occurrence of various notes and note combinations). From Balzano's (1982) perspective, the choice of tonic is not involved in the unique properties of the major scale. Consequently, one might expect sensitivity to the choice of tonic to emerge only after considerable musical exposure. Thus adults should find melodies based on the major scale easier than those based on the Phrygian but infants should show no difference.

Our primary focus on Western scale structure stems from convenience rather than its presumed superiority over alternative musical structures. It is reasonable to expect, then, that scales prominent in other cultures would also embody aspects of *good* form. Some recent evidence bears on this question. Lynch *et al.* (1990) investigated Western infants' (six months of age) and adults' ability to detect mistunings (subtle contour-preserving changes) in melodies based on the Western major and Javanese *pélog* scale. Adults with little training but the usual informal exposure performed better with the major than *pélog* melodies whereas infants performed equivalently on both. These findings imply that, in the early months of life, infants' discrimination of such melodic changes is unaffected by culture-specific experience. This raises the possibility of enhanced processing for the *pélog* scale comparable to that observed for the major scale.

9.6 *GOOD* CHANGES TO *GOOD* PATTERNS

The aforementioned studies of pattern *goodness* indicate that prototypical patterns in Western music embody features that promote detailed encoding in naïve listeners. Thus infants encode not only global information about melodic contour but also local information about interval size. Even if some analytic skills are present in infancy, it is likely that others are schema-based, emerging only after extended exposure. As a result, infants may not be sensitive to all structural aspects of patterns that conform to the rules of a musical system.

Perhaps *good* and *bad* patterns can be distinguished on the basis of universal principles that are evident in infancy. In the context of a *good* pattern, however, or one readily processed by infants, would there be differential sensitivity to changes that conform to Western musical conventions compared with those that violate such conventions? For adults, a comparison melody in which the changed note goes outside the key of a conventionally structured standard melody is easier to detect than a change that remains within the key (Cuddy *et al.* 1979). In fact, the ease with which adults detect such structural violations (i.e. a non-key note) reveals their knowledge of scale structure.

Would infants be capable of processing comparable aspects of key structure? We presented eight-month-olds and young adults with a ten-note conventionally structured melody (in the key of C major: C_4 E_4 G_4 F_4 D_4 G_3 C_4 E_4 D_4 C_4) in transposition (Trainor and Trehub 1992). There were two comparison melodies (see Fig. 9.14). In one, the sixth note (G_3) was raised by a semitone to $A\flat_3$, which was outside the prevailing key and therefore violated Western musical rules. In the second, the same note was raised by four semitones to B_3, which is in the prevailing key. Adults performed as predicted, readily detecting the out-of-key change. In some circumstances, they were at chance levels in detecting the within-key pitch change, even though it was four times as large as the out-of-key change. Infants, on the other hand, performed equivalently and at above chance levels on both pitch changes (see Fig. 9.15). Moreover, they performed significantly better than adults on the within-key change.

Infants' performance is revealing in a number of respects. First, their ability to detect semitone changes in the context of a conventionally structured ten-tone melody increases the generality of our previous findings of enhanced processing (i.e. interval processing) for patterns

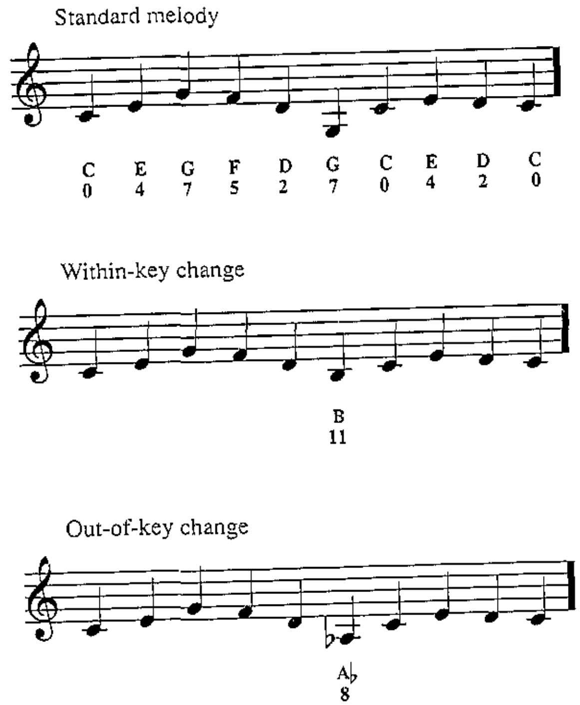

Fig 9.14 The standard ten-note melody (top), the four-semitone within-key change (middle), and the one-semitone out-of-key (bottom) change. (Stimuli from Trainor and Trehub (1992).)

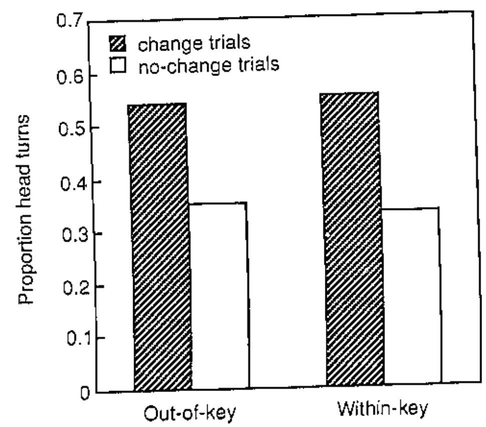

Fig 9.15 Mean proportion of head turns on change and no-change trials for the out-of-key and within-key changes. (Data from Trainor and Trehub (1992).)

based on the major triad. Second, infants' facility in detecting the within-key or structure-conserving change coupled with adults' difficulty on this discrimination implies that infants' performance is culture-independent. Although infants are insensitive to some culturally relevant musical structures, it is difficult to specify their missing knowledge. It is possible, for example, that infants lack tacit knowledge of the notes that belong in a key or scale. As a result, their superior performance for conventional melodies based on the major scale as opposed to unconventional melodies with non-scale notes (Cohen *et al.* 1987; Trainor 1991; Trehub *et al.* 1990) could be due to special properties of the major scale such as the prominence of perfect fifth intervals.

An alternative explanation of this finding is that infants, unlike Western adults, do not process melodies according to their implied harmonic structure. Harmony, the combination of notes sounded simultaneously, or the vertical, as opposed to horizontal (i.e. melodic), structure of music, is a unique feature of Western music (Piston 1969). Thus, one might expect relatively slow emergence of sensitivity to harmonic structure just as rare phonemic distinctions emerge considerably later than those present in many languages (Locke 1983). Even an unaccompanied Western melody is considered to have implied harmonies. For example, the standard melody in Trainor and Trehub (1992), C_4 E_4 G_4 F_4 D_4 G_3 C_4 E_4 D_4 C_4, implies a tonic chord (C_4 E_4 G_4) followed by a dominant chord (F_4 D_4 G_3). Because B_3 (the changed note in one condition) is also part of the dominant chord, this change not only remains within the key of the original melody but also within the implied harmony. A likely possibility, then, is that the performance difference between infants and adults reflects adults' acquired propensity to process the implied harmony of a melody. These findings point to some of the fine tuning or perceptual reorganization that results from extended, informal listening to music. Mapping the timetable of this retuning will be of considerable developmental interest.

9.7 CONCLUSION

At the outset, we posed a number of questions about infants' processing of complex auditory patterns. What answers can we offer on the basis of our findings and those of others? We can say with assurance that perceptual grouping processes such as those described by Bregman (Ch. 2 this volume), Bigand (Ch. 8 this volume), Krumhansl (1992) and others are indeed operative in infancy. Moreover, relational pitch and temporal processing are as characteristic of infants as they are of adults. In particular, relational pitch processing would seem to be a basic and uniquely human disposition, having little to do with acquired knowledge about its utility and more to do with the nature of the human brain. On the other hand, that much vaunted skill, absolute pitch processing, seems to be 'for the birds' (Hulse *et al.* 1990). For reasons that are as yet unclear, extensive childhood exposure to music sometimes results in the addition of 'bird-like' absolute pitch processing to the usual repertoire of human listening strategies (Cohen and Baird 1990).

Pitch contours seem to dominate perception at a time when the infant's world is uncluttered with referential meaning. Thus infants proceed globally, extracting the pitch contours of melodies and spoken utterances and ignoring many of the details within such contours. This is not to say that they are unable to resolve auditory detail. On the contrary, impoverished auditory contexts, such as the typical laboratory experiments with single sounds, reveal impressive phoneme discrimination skills (see Kuhl 1986) and excellent frequency resolution (Olsho 1984) in early infancy.

Our approach to the definition of *good* patterns was empirical rather than theoretical, involving the identification of patterns or pattern elements that are congenial to, or readily processed by, infants. This approach yielded additional perspectives on infants' pattern processing capabilities and provided a means of differentiating universal from culture-specific constraints on musical structure. Specifically, we found that infants could supplement their primary listening strategy of contour processing with a secondary strategy of interval processing, but only in very special circumstances. We know that these circumstances include prototypical melodies from our culture (e.g. melodies based on the major triad) and prototypical contexts (e.g. neighbouring patterns with related keys). What we are likely to learn in the future is that prototypical melodies from foreign musical cultures also promote enhanced processing, engaging both global and analytic strategies. One task that lies ahead is to specify the features associated with such enhanced processing, features that can be considered 'natural' and, therefore, belonging to the set from which musical universals would be drawn.

Our findings cast doubt on some prominent approaches to adult pattern perception. We dispute the view that pattern perception processes have virtually unlimited plasticity, with experience making possible and shaping the perceptual organization of patterns (Espinoza-Varas and Watson 1989). Likewise, we dispute the assumption of a number of twentieth-century composers and music theorists (e.g. Boulez 1971; Forte 1973; Schoenberg 1975) that listeners, given comparable exposure, would perceive or acquire all conceivable types of musical structure as readily as they perceive and acquire conventional musical structure. Not all songs are equivalent as learning stimuli, even for those songbirds that require exposure for song acquisition (Marler 1990). We would argue, instead, that auditory perceptions are initially structured or organized (Trehub 1985; Trehub and Trainor 1990) and that experience leads to alternative organizations including fine tuning to culture-specific circumstances. We would argue, further, that some organizations are more 'natural' or robust than others (Burnham *et al.* 1987; Trehub 1985; Trehub and Unyk 1992) and that the natural organization may prevail throughout life. For example, both infants and adults approach unfamiliar melodies with a pitch contour strategy and, in the case of prototypical melodies, they often add an interval-processing strategy.

Our findings also raise questions about prominent approaches to music perception that involve the internalization of pitch relations or prototypes based on regularities in heard music (e.g. Krumhansl 1990; Jones 1981, 1990). In the absence of data on infant music perception, it may have been reasonable to presume that adults' interval-processing strategy with familiar melodies (Attneave and Olson 1971; Dowling and Fujitani 1971) was attributable to their extensive experience with those melodies and that their superior retention of conventionally structured melodies (e.g. Cuddy *et al.* 1981; Francès 1988) reflected their accumulated knowledge of musical structure. It turns out, however, that such familiar and unfamiliar but conventional melodies have features that promote analytic or interval processing in infancy. From the infant's perspective, however, these melodies are neither familiar nor conventional but they nevertheless elicit interval processing. We would argue, then, that it may be unnecessary for listeners to accumulate information about the frequency of occurrence of structural as opposed to stylistic features of musical patterns, constructing internal models based on such statistical information. In fact, comparable approaches have been rejected as viable strategies for the acquisition of linguistic structures (Chomsky 1968). The more likely alternative is that composers intuitively create or select patterns that build on 'natural' and universal principles of pattern perception (Terhardt 1987). The result is that frequency of occurrence and *naturalness* are often confounded. Learning obviously plays a role in perceptual fine tuning or calibration, in the acquisition of some musical conventions, and in musical production. One important goal for the future is to make further progress in differentiating experience-related from natural listening strategies.

Finally, the listening strategies that we have outlined have implications for language as well as musical development. We have suggested that prelinguistic infants parse the speech stream into pitch contours, which may serve as elementary processing units. Adults seem to cooperate in this endeavor by communicating in a special speech register or style that upgrades contour information at the expense of clear articulation and informative content. The resulting infant-directed speech is a kind of singsong that captures the infant's attention and promotes positive affect (Fernald 1992; Werker and McLeod 1989). In other words, the simple pitch contours of infant-directed speech seem to have special status or perceptual priority for prelinguistic listeners. Later, infants can use contour information to track phrasal and clausal units and also to distinguish utterance types (questions vs. statements). The transition from global processing to analytic or joint global/analytic processing of speech may depend on important cognitive changes that occur late in the first year. Such changes may promote the perceptual reorganization or fine tuning of some speech sound categories at this time (Werker and Lalonde 1988). Just as culture-specific modes of music processing build on pattern processing predispositions, so mature speech processing seems to build on inherent speech discrimination skills that are evident in infancy (Best *et al.* 1988; Burnham *et al.* 1987; Werker and Lalonde 1988). On the whole, however, pitch-contour processing seems to dominate infants' early perception of connected speech and music. One consequence for production is that native-like intonation contours emerge well before the first words (de Boysson-Bardies *et al.* 1984; Crystal 1973) just as appropriate musical contours predate correct intervals (Davidson *et al.* 1981; Kelley and Sutton-Smith 1987).

It is unclear, however, whether there are speech parallels to the enhanced processing of prototypical musical patterns. In this regard, there is suggestive evidence that *good* or prototypical exemplars of vowels have a more robust representation for infant listeners than do *poor* or less typical exemplars (Grieser and Kuhl 1989; Kuhl 1991). The intonation contours that are prevalent in infant-directed speech across cultures may also function as prototypes for infant listeners and their caretakers but there has been no direct investigation of this question. One perceptual consequence of prototypical status would be enhanced differentiation of such contours from others.

The findings of enhanced processing for major triad melodies and near-key relations raise the possibility that the perfect fifth interval, which approximates a 2:3 frequency ratio, is inherently *good*. Not only does the perfect fifth interval figure prominently in musical sounds, both simultaneous and successive, it also occurs commonly in the simultaneous components (partials or overtones) of complex sounds in the natural environment, including vowel sounds. The special emphasis accorded to vowels in infant-directed speech (i.e. they are greatly extended in duration) may highlight the perfect fifth interval (simultaneously sounded) and contribute to the special status of such speech. Whether the special qualities derive from the perfect fifth interval in particular or intervals that approximate simple ratios in general remains to be determined. Perhaps the relative prominence of the perfect fifth interval affects adults' ratings of the *goodness* of vowel sounds (Grieser and Kuhl 1989; Kuhl 1991) but this has not been evaluated to date. In any case, the cross-cultural prevalence of simple ratios in informal vocal music (Kolinski 1967) is suggestive of their importance.

In the early months of life when the processing of speech and musical sounds is largely (but not entirely) culture-independent, it would not be surprising for both domains to be dominated by common processing tendencies. We have suggested repeatedly that pitch-contour processing is prominent among these processing dispositions, with infants extracting the pitch contours of simple speech and musical passages.

It is also possible that some contours are inherently *good* for infant listeners and that these contours are prominent in speech and song to infants. One example of this is the prevalence of descending contours in soothing infant-directed speech (Fernald and Simon 1984) and music (Unyk *et al.* 1992). Also of interest is the prominence of descending contours in infant vocalization (Fox 1990). Simplicity (e.g. few directional changes in pitch) may be one factor that determines the relative *goodness* of a contour and its resultant ease of encoding. Indeed, simple contours are common in maternal speech (Fernald 1984) as they are in lullabies (Unyk *et al.* 1992) and in infant vocalizations (Delack and Fowlow 1978; Fox 1990). According to one music-theoretic framework (Narmour 1990), adult listeners, regardless of their experience and cultural origins, have certain common expectations about pitch movement within a melody. Preliminary support for this notion has been obtained with American and Chinese listeners (Schellenberg and Krumhansl, personal communication). If such expectations have an innate basis, as Narmour (1990) contends, they might be present in infancy, contributing to the essence of *good* patterns.

The pattern of intervals may also be implicated in the *goodness* of speech and musical patterns. For example, the major triad is asymmetric (a four-semitone interval followed by a three-semitone interval), being derived from a scale with unequal interval structure. The feature of asymmetry in scale systems from which melodic sequences are derived is considered to facilitate the functional differentiation of component notes (Balzano 1982; Butler and Brown 1984; Krumhansl 1987). It is unclear, however, whether comparable asymmetries characterize the intonation patterns of infant-directed speech. Just as there are rules for combining notes into acceptable melodic sequences, so are there rules for combining speech sounds into acceptable utterances. There are suggestions, moreover, that some sequences of speech sounds are more learnable than others (Locke 1990). Perhaps this is also the case for intonation patterns.

Regardless of the ultimate list of common features that will emerge from future studies of speech and musical processing, it is clear that some common features are present. Why might this be the case? In the prelinguistic period, speech directed to infants is not meaningful in the way that it is later on. Rather, the meanings are affective in nature and are conveyed principally through intonation and 'tone of voice' (Fernald 1989; Papoušek and Papoušek 1981). Jusczyk and Bertoncini (1988) have proposed that development in the domain of speech perception is facilitated by innately guided learning such as that described by Gould and Marler (1987) for the acquisition of important species-specific abilities. This process is characterized by considerable selectivity in responsiveness to signals and rapid learning on the basis of limited culture-specific experience. Innately guided learning contrasts with incremental learning, which implicates a flexible and initially unselective system that gradually achieves appropriate organization of the input on the basis of its distributional properties (Mehler and Dupoux, 1990). Early differentiation of the mother's voice from other voices (DeCasper and Fifer 1980; Mehler *et al.* 1978), of the native 'language-to-be' from other languages (Mehler *et al.* 1988), and of infant-directed from adult-directed speech (e.g. Cooper and Aslin 1990; Fernald 1985) are possible examples of such innately guided learning.

The music perception skills that we have uncovered in infants may also reflect innately guided learning, perhaps the same or similar processes that promote the development of speech perception skills (e.g. Werker and Tees 1984) and the appearance of language-specific aspects of babbling (de Boysson-Bardies *et al.* 1989) in prelinguistic infants. It is possible, then, that precocious abilities in both domains stem from common origins. If so, we would expect the processing of speech and music to diverge once infants are ready to assign referential meaning to speech. We know that after the hypothesized separation, the development of speech perception and production skills continues at a rapid pace. It is unclear, however, whether this is also the case for music perception and production. Accordingly, it will be of particular interest to determine whether the precocity of music processing in infancy is dependent on its initial ties to speech. If further progress in music perception and production is considerably slower than in the language domain but comparable to other biologically non-significant domains, this would suggest accidental convergence of speech and music in early infancy. If, on the other hand, music perception continues to exhibit a developmental pace characteristic of innately guided learning, this would revive interest in its biological significance (Granit 1977; Lerdahl and Jackendoff 1983).

ACKNOWLEDGEMENTS

The preparation of this chapter was assisted by grants to the senior author from the Natural Sciences and Engineering Research Council of Canada and the University of Toronto.

REFERENCES

Aslin, R. N. and Smith, L. B. (1988). Perceptual development. *Annual Review of Psychology*, **39**, 435–73.

Attneave, F. and Olson, R. K. (1971). Pitch as a medium: a new approach to psychophysical scaling. *American Journal of Psychology*, **84**, 147–66.
Balzano, G. J. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal*, **4**, 66–84.
Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In *Music, mind, and brain: the neuropsychology of music* (ed. M. Clynes), pp. 321–51. Plenum, New York.
Bartlett, J. C. and Dowling, W. J. (1980). Recognition of transposed melodies: a key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 501–15.
Bebey, F. (1969). *African music: a people's art*. Lawrence Hill, Westpoint.
Beebe, B., Feldstein, S., Jaffe, J., Mays, K., and Alson, D. (1985). Interpersonal timing: the application of an adult dialogue model to mother-infant vocal and kinesic interactions. In *Social perception in infants* (ed. T. M. Field and N. A. Fox), pp. 217–47. Ablex, Norwood, NJ.
Best, C. T., McRoberts, G. W., and Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 45–60.
Bharucha, J. J. (1984). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology*, **16**, 485–518.
Bharucha, J. J. (1987). Music cognition and perceptual facilitation: a connectionist framework. *Music Perception*, **5**, 1–30.
Bharucha, J. J. and Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: an asymmetry in discrimination. *Perception and Psychophysics*, **40**, 137–41.
Bharucha, J. J., Olney, K. L., and Schnurr, P. P. (1985). Detection of coherence-disrupting and coherence-conferring alterations in text. *Memory and Cognition*, **13**, 573–8.
Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, **6**, 145–238.
Boulez, P. (1971). *Boulez on music today*. (transl. S. Bradshaw and R. R. Bennett). Faber and Faber, London.
Brakeley, T. C. (1950). Lullaby. In *Standard dictionary of folklore, mythology, and legend*, pp. 653–4. Funk and Wagnalls, New York.
Bregman, A. S. (1978). The formation of auditory streams. In *Attention and performance. Vol. VII* (ed. J. Requin), pp. 63–75. Erlbaum, Hillsdale, NJ.
Bregman, A. S. (1981). Asking the 'what for' question in auditory perception. In *Perceptual organization* (ed. M. Kubovy and J. R. Pomerantz), pp. 99–118. Erlbaum, Hillsdale, NJ.
Bregman, A. S. (1990). *Auditory scene analysis: the perceptual organization of sound*. MIT, Cambridge, MA.
Bregman, A. S. and Campbell, J. (1971). Primary auditory stream segregation and the perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, **89**, 244–9.
Brown, M. J. E. (1980). Lullaby. In *The new Grove dictionary of music and musicians*, pp. 313–4. Macmillan, London.
Burnham, D. K., Earnshaw, L. J., and Quinn, M. C. (1987). The development of the categorical identification of speech. In *Perceptual development in early infancy: problems and issues* (ed. B. E. McKenzie and R. J. Day), pp. 237–75. Erlbaum, Hillsdale, NJ.

Burnham, D. K., Earnshaw, L. J., and Clark, J. E. (1991). Development of categorical identification of native and non-native bilabial stops: infants, children, and adults. *Journal of Child Language*, **18**, 231–60.
Burns, E. M. and Ward, W. D. (1982). Intervals, scales, and tuning. In *The psychology of music* (ed. D. Deutsch), pp. 241–69. Academic, Orlando, FL.
Butler, D. (1989). Describing the perception of tonality in music. A critique of the tonal hierarchy theory and proposal for a theory of intervallic rivalry. *Music Perception*, **6**, 219–42.
Butler, D. and Brown, H. (1984). Tonal structure versus function: studies of the recognition of harmonic motion. *Music Perception*, **2**, 6–24.
Cass-Beggs, B. and Cass-Beggs, M. (1969). *Folk lullabies*. Oak Publications, New York.
Chang, H. W. and Trehub, S. E. (1977*a*). Infants' perception of temporal grouping in auditory patterns. *Child Development*, **48**, 1666–70.
Chang, H. W. and Trehub, S. E. (1977*b*). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, **24**, 324–31.
Chomsky, N. (1968). *Language and mind*. Harcourt Brace Jovanovich, New York.
Clifton, R. K. (1992). The development of spatial hearing in human infants. In *Developmental psychoacoustics* (ed. L. A. Werner and E. W. Rubel) pp. 135–57. American Psychological Association, Washington, DC.
Cohen, A. J. and Baird, R. (1990). Acquisition of absolute pitch: the question of critical periods. *Psychomusicology*, **9**, 31–7.
Cohen, A. J., Thorpe, L. A., and Trehub, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology*, **41**, 33–47.
Comrie, B. (1981). *Language universals and linguistic typology*. Blackwell, Oxford.
Cong-Huyen-Ton-Nu, N.-T. (1979). The functions of folk songs in Vietnam. In *The performing arts: music and dance* (ed. J. Blacking and J. W. Kealiinohomoku), pp. 141–51. Mouton, The Hague.
Cooper, R. B. and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, **61**, 1584–95.
Crystal, D. (1973). Non-segmental phonology in language acquisition: a review of the issues. *Lingua*, **32**, 1–45.
Cuddy, L. L., Cohen, A. J., and Miller, J. (1979). Melody recognition: the experimental application of musical rules. *Canadian Journal of Psychology*, **33**, 148–56.
Cuddy, L. L., Cohen, A. J., and Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 869–83.
D'Amato, M. R. (1988). A search for tonal pattern perception in cebus monkeys: why monkeys can't hum a tune. *Music Perception*, **5**, 453–80.
D'Amato, M. R. and Salmon, D. P. (1984). Processing of complex auditory stimuli (tunes) by rats and monkeys (*Cebus apella*). *Animal Learning and Behavior*, **12**, 184–94.
Davidson, L., McKernon, P., and Gardner, H. (1981). The acquisition of song: a developmental approach. In *Documentary report of the Ann Arbor Symposium* (pp. 301–15). Music Educators National Conference, Reston, Virginia.
de Boysson-Bardies, B., Sagart, L., and Durand, C. (1984). Discernible differences

in the babbling of infants according to target language. *Journal of Child Language*, **11**, 1–15.
de Boysson-Bardies, B., Hallé, P., Sagart, L., and Durand, C. (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language*, **16**, 1–17.
DeCasper, A. J. and Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, **208**, 1174–6.
Delack, J. B. and Fowlow, P. J. (1978). The ontogenesis of differential vocalization: development of prosodic contrastivity during the first year of life. In *The development of communication* (ed. M. Waterson and C. Snow), pp. 93–110. Wiley, Chichester.
Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development*, **5**, 261–76.
Demany, L. and Armand, F. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America*, **76**, 57–66.
Demany, L., McKenzie, B., and Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, **266**, 718–19.
Deutsch, D. (1982). Grouping mechanisms in music. In *The psychology of music* (ed. D. Deutsch), pp. 99–134. Academic, New York.
Diehl, R. L., Buchwald McCusker, S., and Chapman, L. S. (1981). Perceiving vowels in isolation and in consonantal context. *Journal of the Acoustical Society of America*, **69**, 239–48.
Dooling, R. J., Brown, S. D., Park, T. J., Okanoya, K., and Soli, S. D. (1987). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): I. Pure tones. *Journal of Comparative Psychology*, **101**, 139–49.
Dooling, R. J., Brown, S. D., Park, T. J., and Okanoya, K. (1990). Natural perceptual categories for vocal signals in budgerigars (*Melopsittacus undulatus*). In *Comparative perception. Vol. 2: Complex signals* (ed. W. C. Stebbins and M. A. Berkley), pp. 345–74. Wiley, New York.
Dorman, M. F. and Raphael, L. J. (1980). Distribution of acoustic cues for stop consonant place of articulation in VCV syllables. *Journal of the Acoustical Society of America*, **67**, 1333–5.
Dowling, W. J. (1978). Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, **85**, 341–54.
Dowling, W. J. and Fujitani, D. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, **49**, 524–31.
Dowling, W. J. and Harwood , D. L. (1986). *Music cognition*. Academic, Orlando, FL.
Eilers, R. E., Wilson, W. R., and Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research*, **20**, 766–80.
Eimas, P. D. and Miller, J. L. (1980). Contextual effects in infant speech perception. *Science*, **209**, 1140–1.
Espinoza-Varas, B. and Watson, C. S. (1989). Perception of complex auditory patterns by humans. In *The comparative psychology of audition: perceiving complex sounds* (ed. R. J. Dooling and S. H. Hulse), pp. 67–94. Erlbaum, Hillsdale, NJ.
Fernald, A. (1984). The perceptual and affective salience of mothers' speech to

infants. In *The origins and growth of communication* (ed. L. Feagans, C. Garvey, and R. Golinkoff), pp. 5–29. Ablex, Norwood, NJ.
Fernald, A. (1985). Four-month-old infants prefer to listen to Motherese. *Infant Behavior and Development*, **8**, 181–95.
Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Development*, **60**, 1497–510.
Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In *Nonverbal vocal communication: comparative and developmental approaches* (ed. H. Papoušek, V. Jurgens, and M. Papoušek), pp. 262–82. Cambridge University Press.
Fernald, A. and Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese. *Infant Behavior and Development*, **10**, 279–93.
Fernald, A. and Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, **20**, 104–13.
Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., de Boysson-Bardies, B., and Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, **16**, 477–501.
Fetterman, J. G. (1987). Same–different comparison of duration. *Animal Learning and Behavior*, **15**, 403–11.
Fitzgibbons, P. J., Pollatsek, A., and Thomas, I. B. (1974). Detection of temporal gaps within and between perceptual tonal groups. *Perception and Psychophysics*, **16**, 522–8.
Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology*. MIT, Cambridge, MA.
Forte, A. (1973). *The structure of atonal music*. Yale University Press, New Haven, CT.
Fox, D. B. (1990). An analysis of the pitch characteristics of infant vocalizations. *Psychomusicology*, **9**, 21–30.
Fraisse, P. (1982). Rhythm and tempo. In *The psychology of music* (ed. D. Deutsch), pp. 149–80. Academic, New York.
Francès, R. (1988). *The perception of music*. Erlbaum, Hillsdale, NJ (transl. W. J. Dowling from *La perception de la musique*, 1958; 2nd. edn. 1984. Vrin, Paris.)
Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist*, **58**, 34–42.
Garner, W. R. (1974). *The processing of information and structure*. Erlbaum, Hillsdale, NJ.
Gould, J. L and Marler, P. (1987). Learning by instinct. *Scientific American*, **256**, 74–85.
Granit, R. (1977). *The purposive brain*. MIT Press, Cambridge, MA.
Grieser, D. L. and Kuhl, P. K. (1989). Categorization of speech by infants: support for speech-sound prototypes. *Developmental Psychology*, **25**, 577–88.
Handel, S. (1989). *Listening: an introduction to the perception of auditory events*. MIT, Cambridge, MA.
Hargreaves, D. M. (1986). *The developmental psychology of music*. Cambridge University Press.
Harwood, D. L. (1976). Universals in music: a perspective from cognitive psychology. *Ethnomusicology*, **20**, 521–34.

Helmholtz, H. L. F. (1954). *On the sensations of tone as a physiological basis for the theory of music* (transl. A. J. Ellis from German 4th edn 1877), Dover, New York (reprinted from English 2nd edn, 1885).
Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., and Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, **26**, 269–86.
Hockett, C. F. (1963). The problem of universals in language. In *Universals in language* (ed. J. H. Greenberg), pp. 1–29. MIT Press, Cambridge, MA.
Howell, P., Cross, I., and West, R. (ed.) (1985). *Musical structure and cognition*. Academic, London.
Hulse, S. H. and Cynx, J. (1986). Interval and contour in serial pitch perception by a passerine bird, the European starling (*Sturnus vulgaris*). *Journal of Comparative Psychology*, **100**, 215–28.
Hulse, S. H., Humpal, J., and Cynx, J. (1984). Discrimination and generalization of rhythmic and arhythmic sound patterns by European starlings (*Sturnus vulgaris*). *Music Perception*, **1**, 442–64.
Hulse, S. H., Page, S. C., and Braaten, R. F. (1990). An integrative approach to auditory perception by songbirds. In *Comparative perception. Vol. 2: Complex signals* (ed. W. C. Stebbins and M. A. Berkley), pp. 3–34. Wiley, New York.
Jaffe, J., Anderson, S., and Stern, D. (1979). Conversational rhythms. In *Psycholinguistic research* (ed. D. Aronson and R. Rieber). Erlbaum, Hillsdale, NJ.
Jakobson, R. (1968). *Child language, aphasia, and phonological universals*. Mouton, The Hague.
Jones, M. R. (1981). Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology*, **1**, 14–31.
Jones, M. R. (1982). Music as a stimulus for psychological motion: Part II. An expectancy model. *Psychomusicology*, **2**, 1–13.
Jones, M. R. (1990). Learning and the development of expectancies: an interactionist approach. *Psychomusicology*, **9**, 193–228.
Jusczyk, P. W. and Bertoncini, J. (1988). Viewing the development of speech perception as an innately guided learning process. *Language and Speech*, **31**, 217–38.
Kelley, L. and Sutton-Smith, B. (1987). A study of infant musical productivity. In *Music and child development* (ed. J. C. Peery, I. Weiss Peery, and T. W. Draper), pp. 35–53. Springer, New York.
Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., and Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, **16**, 66–8.
Kolinski, M. (1967). Recent trends in ethnomusicology. *Ethnomusicology*, **11**, 1–24.
Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception*, **1**, 28–62.
Krumhansl, C. L. (1987). General properties of musical pitch systems: some psychological considerations. In *Harmony and tonality* (ed. J. Sundberg), Publication No. 54. Royal Swedish Academy of Music, Stockholm, Sweden.
Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press, New York.
Krumhansl, C. L. (1992). Grouping processes in infants' music perception. In *Grouping in music* (ed. J. Sundberg, L. Nord, and R. Carlson). Royal Swedish Academy of Music, Stockholm, Sweden. (In press.)

Krumhansl, C. L. and Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science*, **1**, 70–3.
Krumhansl, C. L. and Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition*, **10**, 243–51.
Krumhansl, C. L., Bharucha, J. J., and Castellano, M. A. (1982*a*). Key distance effects on perceived harmonic structure in music. *Perception and Psychophysics*, **32**, 579–94.
Krumhansl, C. L., Bharucha, J. J., and Kessler, E. J. (1982*b*). Perceived harmonic structure of chords in three related keys. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 24–36.
Kuhl, P. K. (1985). Methods in the study of infant speech perception. In *Measurement of audition and vision in the first year of postnatal life: a methodological overview* (ed. G. Gottlieb and N. A. Krasnegor), pp. 223–51. Ablex, Norwood, NJ.
Kuhl, P. K. (1986). The special-mechanisms debate in speech research: categorization tests on animals and infants. In *Categorical perception* (ed. S. Harnad), pp. 355–86. Cambridge University Press.
Kuhl, P. K. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, **50**, 93–107.
Kuhl, P. K. and Miller, J. D. (1975). Speech perception by the chinchilla: voiced–voiceless distinction in alveolar plosive consonants. *Science*, **190**, 69–72.
Kuhl, P. K. and Padden, D. M. (1982). Enhanced discriminability at the phonetic boundaries for voicing feature in macaques. *Perception and Psychophysics*, **32**, 542–50.
Langlois, J. H. and Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, **1**, 115–21.
Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., and Jenkins, V. Y. (1987). Infant preferences for attractive faces: rudiments of a stereotype? *Developmental Psychology*, **23**, 363–9.
Leek, M. R. (1987). Directed attention in complex sound perception. In *Auditory processing of complex sounds* (ed. W. A. Yost and C. S. Watson), pp. 278–88. Erlbaum, Hillsdale, NJ.
Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music*. MIT, Cambridge, MA.
Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, **74**, 431–61.
Locke, J. L. (1983). *Phonological acquisition and change*. Academic, New York.
Locke, J. L. (1990). Structure and stimulation in the ontogeny of spoken language. *Developmental Psychobiology*, **23**, 621–43.
Lynch, M. P., Eilers, R. E., Oller, D. K., and Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science*, **1**, 272–6.
McCosker, S. S. (1974). *The lullabies of the San Blas Cuna Indians of Panama*. Gothenburg Ethnographical Museum, Gothenburg.
McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746–8.
Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press.
Marler, P. (1990). Innate learning preferences: signals for communication. *Developmental Psychobiology*, **23**, 557–68.

Mathews, M. V, Pierce, J. R., and Roberts, L. A. (1987). Harmony and new scales. In *Harmony and tonality* (ed. J. Sundberg), Publication No. 54. Royal Swedish Academy of Music, Stockholm, Sweden.
Mehler, J. and Dupoux, E. (1990). *Naître humain*. Odile Jacob, Paris.
Mehler, J., Bertoncini, J., Barrière, M., and Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception*, **7**, 491–7.
Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, **29**, 143–78.
Mervis, C. B. and Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, **32**, 89–115.
Miller, J. L. and Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception and Psychophysics*, **25**, 457–65.
Miller, J. L. and Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception and Psychophysics*, **46**, 505–12.
Moorhead, G. E. and Pond, D. (1978). *Music of young children*. Pillsbury Foundation, Santa Barbara, CA.
Morgan, J. L., Meier, R. P., and Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, **19**, 498–550.
Morrongiello, B. A. (1984). Auditory temporal pattern perception in 6- and 12-month-old infants. *Developmental Psychology*, **20**, 441–8.
Morrongiello, B. A. (1988). The development of auditory pattern perception skills. In *Advances in infancy research*, Vol. 5 (ed. C. Rovee-Collier), pp. 135–72. Ablex, Norwood, NJ.
Muir, D. W., Clifton, R. K., and Clarkson, M. G. (1989). The development of a human auditory localization response: a U-shaped function. *Canadian Journal of Psychology*, **43**, 199–216.
Narmour, E. (1990). *The analysis and cognition of basic melodic structures: the implication–realization model*. University of Chicago Press.
Neisser, U. (1967). *Cognitive psychology*. Appleton-Century-Crofts, New York.
Olsho, L. W. (1984). Infant frequency discrimination. *Infant Behavior and Development*, **7**, 27–35.
Papoušek, H. and Papoušek, M. (1987). Intuitive parenting: a dialectic counterpart to the infant's integrative competence. In *Handbook of infant development*, 2nd edn (ed. J. Osofsky), pp. 669–720. Wiley, New York.
Papoušek, M. and Papoušek, H. (1981). Musical elements in the infant's vocalization: their significance for communication, cognition, and creativity. In *Advances in infancy research*, Vol. 1 (ed. L. P. Lipsitt), pp. 163–224. Ablex, Norwood, NJ.
Papoušek, M., Papoušek, H., and Bornstein, M. H. (1985). The naturalistic vocal environment of young infants: on the significance of homogeneity and variability in parental speech. In *Social perception in infants* (ed. T. M. Field and N. A. Fox), pp. 269–97. Ablex, Norwood, NJ.
Piston, W. (1969). *Harmony*. Norton, New York.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music*, **18**, 60–72.
Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, **7**, 532–47.
Sakata, H. L. (1987). Hazara women in Afghanistan: innovators and preservers of a musical tradition. In *Women and music in cross-cultural perspective* (ed. E. Koskoff), pp. 85–95. Greenwood, Westport, CT.
Sands, K. M. and Sekaquaptewa, E. (1978). Four Hopi lullabies. A study in method and meaning. *American Indian Quarterly*, **4**, 195–210.
Schenker, H. (1954). *Harmony* (ed. O. Jones), (transl. E. M. Borgese from 1906 German edn). MIT Press, Cambridge, MA.
Schoenberg, A. (1975). *Style and idea* (ed. L. Stein), (transl. L. Black). Faber and Faber, London.
Slobin, M. and Titon, J. T. (1984). The music-culture as a world of music. In *Worlds of music* (ed. J. T. Titon, J. T. Koetting, D. P. McAllester, D. B. Beck, and M. Slobin), pp. 1–11. Schirmer, New York.
Sloboda, J. A. (1985). *The musical mind: the cognitive psychology of music*. Clarendon, Oxford.
Stern, D. N., Spieker, S., and MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, **18**, 727–35.
Stern, D. N., Spieker, S., Barnett, R. K., and MacKain, K. (1983). The prosody of maternal speech: infant age and context related changes. *Journal of Child Language*, **10**, 1–15.
Stevens, K. N. and Keyser, S. J. (1989). Primary features and their enhancement in consonants. *Language*, **65**, 81–106.
Studdert-Kennedy, M. (1975). From continuous signal to discrete message: syllable to phoneme. In *The role of speech in language* (eds. J. F. Kavanaugh and J. E. Cutting), pp. 113–25. MIT Press, Cambridge, MA.
Summerfield, A. Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 1074–95.
Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, **55**, 1061–9.
Terhardt, E. (1987). Gestalt principles and music perception. In *Auditory processing of complex sounds* (ed. W. A. Yost and C. S. Watson), pp. 157–66. Erlbaum, Hillsdale, NJ.
Thorpe, L. A. (1985). *Auditory-temporal organization: developmental perspectives*. Ph.D. thesis, University of Toronto.
Thorpe, L. A. and Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, **25**, 122–7.
Thorpe, L. A., Trehub, S. E., Morrongiello, B. A., and Bull, D. (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology*, **24**, 484–91.
Trainor, L. J. (1991). *The origins of musical pattern perception: a comparison of infants' and adults' processing of melody*. Ph.D. thesis, University of Toronto.
Trainor, L. J. and Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western tonal structure. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 394–402.

Trainor, L. J. and Trehub, S. E. (1993). Musical context effects in infants and adults: key distance. *Journal of Experimental Psychology: Human Perception and Performance*. (In press.)

Trehub, S. E. (1985). Auditory pattern perception in infancy. In *Auditory development in infancy* (ed. S. E. Trehub and B. A. Schneider), pp. 183–95. Plenum, New York.

Trehub, S. E. (1989). Infants' perception of musical sequences: implications for language acquisition. *Journal of Speech–Language Pathology and Audiology*, **13**, 3–11.

Trehub, S. E. (1990). The perception of musical patterns by human infants: the provision of similar patterns by their parents. In *Comparative perception. Vol. 1: Basic mechanisms* (ed. M. Berkley and W. C. Stebbins), pp. 429–59. Wiley, New York.

Trehub, S. E. and Thorpe, L. A. (1989). Infants' perception of rhythm. Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology*, **43**, 217–29.

Trehub, S. E. and Trainor, L. J. (1990). Rules for listening in infancy. In *The development of attention: research and theory* (ed. J. Enns), pp. 87–119. Elsevier, Amsterdam.

Trehub, S. E. and Unyk, A. M. (1992). Music prototypes in developmental perspective. *Psychomusicology*, **10**, 31–45.

Trehub, S. E., Bull, D. and Thorpe, L. A. (1984). Infants' perception of melodies: the role of melodic contour. *Child Development*, **55**, 821–30.

Trehub, S. E., Thorpe, L. A. and Morrongiello, B. A. (1985). Infants' perception of melodies: changes in a single tone. *Infant Behavior and Development*, **8**, 213–23.

Trehub, S. E., Thorpe, L. A. and Morrongiello, B. A. (1987). Organizational processes in infants' perception of auditory patterns. *Child Development*, **58**, 741–9.

Trehub, S. E., Thorpe, L. A., and Trainor, L. J. (1990). Infants' perception of *good* and *bad* melodies. *Psychomusicology*, **9**, 5–19.

Trehub, S. E., Unyk, A. M., and Trainor, L. J. (1992). Adults identify infant-directed music across cultures. *Infant Behavior and Development* (In press.)

Trehub, S. E., Trainor, L. J., and Unyk, A. M. (1993*a*). Music and speech processing in the first year of life. *Advances in Child Behavior and Development*, **24**. (In press.)

Trehub, S. E., Unyk, A. M., and Trainor, L. J. (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development* (In press.)

Unyk, A. M., Trehub, S. E., Trainor, L. J., and Schellenberg, E. G. (1992). Lullabies and simplicity: a cross-cultural perspective. *Psychology of Music*, **20**, 15–28.

Walker, R. (1987). Musical perspectives on psychological research and music education. *Psychology of Music*, **15**, 167–86.

Walker, R. (1990). *Musical beliefs: psychoacoustic, mythical , and educational perspectives*. Teacher's College Press, New York.

Watkins, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception and Psychophysics*, **37**, 275–88.

Watson, C. S. (1987). Uncertainty, informational masking, and the capacity of immediate auditory memory. In *Auditory processing of complex sounds* (ed. W. A. Yost and C. S. Watson), pp. 267–77. Erlbaum, Hillsdale, NJ.

Werker, J. F. and Lalonde, C. E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental Psychology*, **24**, 672–83.

Werker, J. F. and McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, **43**, 230–46.

Werker, J. F. and Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, **7**, 49–63.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology*, **14**, 1–66.

Younger, B. and Gottlieb, S. (1988). Development of categorization skills: changes in the nature or structure of infant form categories. *Developmental Psychology*, **24**, 611–19

Zenatti, A. (1969). *Le développement génétique de la perception musicale* (Monographies Françaises de Psychologie, No. 17). Centre National de la Recherche Scientifique, Paris.